Internal-combustion engines

Thermal engines

Operating principles and concept

Internal-combustion engines are classified as thermal engines. The essential feature of a thermal engine is cycle direction, which is characterized by the output of work.

Contrasting with thermal engines are heat pumps, also called refrigerating machines, which are characterized by cycle direction in the opposite direction and require drive power in order to operate.

The operating principle of thermal engines is always the same. A working medium is compressed, whereupon energy is supplied while the medium is compressed with a corresponding further increase in pressure. This is followed by an expansion with power output. In open cycles the working medium which has performed the work is discharged. In closed cycles the initial state must be re-established by cooling the working medium before compression is restarted in open and closed cycles.

Table 1: Characteristics and functioning principles of important thermal engines

Thermal engine	Steam circuit	Stirling	Steam engine	Gas turbine	Jet engine	Reciprocating piston engine	Wankel rotary engine
Thermodynamic reference cycle	Rankine	Ericson	Steam cycle	Joule	Joule	Seiliger (Otto/ Diesel)	Seiliger
Typical working medium	H_2O, ethanol	Air, helium	H_2O	Air	Air	Air, air/ fuel mixture	Air, air/ fuel mixture
Cycle direction	Closed		Open				
Energy input	From outside by heat transfer			From inside			
Thermodynamic energy input	Stationary	Non-stationary	Non-stationary	Stationary	Stationary	Non-stationary	Non-stationary
Typical energy sources, fuels	Coal, fuels, uranium	Any heat source	Coal, fuels	Methane, ethane, propane, butane	Kerosine	Gasoline, Diesel fuel,	Gasoline
Transmission of work performed	Turbine	Reciprocating piston, rotary piston (rotor)	Piston	Turbine	Linear momentum	Reciprocating piston	Rotary piston (rotor)
Typical maximum pressure	50 bar	3 bar (air)	50 bar	40 bar	40 bar	200 bar	60 bar
Typical minimum pressure	0.05 bar	1 bar	1 bar	1 bar	1 bar	1 bar	1 bar
Typical max. efficiency	40%	30%	~25%	40%	40%	~42% / ~45%	~30 to 35%
Operating concept	Compression, Heating, Evaporation, Superheating, Expansion, Condensing	Compression, Heating, Expansion, Cooling	Compression, Heating, Evaporation, Superheating, Expansion, Exhaust	Induction, Compression, Combustion, Expansion		Induction, Compression, Combustion, Expansion, Exhaust	

Many thermal engines are characterized by an energy input resulting from a combustion process (Table 1). During combustion the energy chemically bound in the fuel is supplied as reaction heat to the cycle. Here compounds containing carbon and hydrogen oxidize with oxygen, which is why typically ambient air with approximately 21 % volume of oxygen makes up a significant part of the working medium.

Cycle direction

The factor that is crucial to cycle direction is the energy input. Here, a distinction is initially made between stationary (or continuous) energy input and non-stationary (or cyclical) energy input. What all piston engines including the Stirling engine have in common is a non-stationary energy input which only ever occurs close to compression dead center with minimum cylinder volume.

A particular feature of all open cycles is an internal energy input which is achieved by the addition and combustion of fuel. In contrast, closed cycles require an energy input via heat exchangers. Here there is no direct contact between the working medium and combustion processes, with the exception of heat conduction. Unique in this context is the steam engine, which vaporizes the working medium by means of a heat flow from an external source and then supplies it to the piston engine.

Another particular feature of the various thermal engines is the different energy sources which are used. Three different energy sources can be used: solids, liquids and gases. The main advantage of thermal engines which operate with an open cycle and an internal energy input is that they do not require any heat exchangers for cycle direction and are therefore compact in design. This advantage can be amplified by using liquid fuels with a high energy density. Gas engines, for example when deployed in passenger cars and commercial vehicles, are also becoming more attractive (running costs, good fuel consumption). The internal-combustion engine can thus be derived from a selection of thermal engines.

Efficiency

The internal-combustion engine is characterized by open cycle direction with internal combustion. Non-stationary operation facilitates the intake and compression of the working medium at mass-averaged peak temperatures above 2,500 K and averaged peak pressures above 200 bar with a very good maximum efficiency above 40 %.

Stationary cycles, being limited by the properties of the materials, do not achieve mass-averaged peak pressures and temperatures of this magnitude, but instead achieve maximum local peak temperatures around 2,500 K. Gas turbines therefore operate at a lower level of efficiency. The closed cycle of the steam circuit achieves a higher level of efficiency than gas turbines with a likewise moderate pressure of around 50 bar. This is achieved by a significant lowering of the low-pressure level. The other thermal engines are much lower in terms of their maximum efficiency.

Reciprocating-piston engine

The internal-combustion engine, in its reciprocating-piston-engine variation, is the primary thermal engines used for mobile applications. A wide range of fuels is conceivable in principle, but diesel and gasoline are still the main energy sources used.

Real cycles

A cycle is a thermodynamic process which has identical initial and final states (see Thermodynamics). The cycle usually passes through several changes of state in such a way that with a thermal engine work is drawn from the cycle. Here the working medium of the cycle undergoes thermodynamic changes of state.

Ideal reference cycles are suitable for demonstrating basic correlations. In the case of new, as yet unknown engines, they help to provide an overview of their method of operation and efficiencies. However, a real cycle calculation is required for detailed analyses.

Real cycles are set in proportion to the ideal cycles. This means that the heat capacities are considered for example depending on temperature or pressure.

The chemically altered composition of the smoke gases is also approximated in the substance variables in order to take into consideration the changed physical properties based on combustion. In particular, what is assumed is not an adiabatic change of state, but rather at least one polytropic curve with an exponent adapted to the heat loss, or even the heat loss to the wall is used, e.g. with the Woschni approach from the Reynolds similarity theory [1].

The charge cycle is calculated with its dissipation losses (flow losses, real flow cross-sections, see Fluid mechanics) and thus the residual exhaust gas is also taken into account. Furthermore, empirical approaches are normally used for friction and the fuel calorific value is calculated depending on the air ratio. Ultimately, in particular the heat addition (combustion process, heating process) and the heat dissipation (heat transfer) are modeled in detail.

One way to quickly assess real cycles is to describe them using an efficiency chain. Here the real cycles are successively depicted by taking individual characteristic quantities into account.

Overall efficiency

The overall efficiency or effective efficiency η_{eff} sets the effectively available power P_{eff} in proportion to the energy flow $\dot{Q}_{add} = \dot{m}_B H_u$ which was added by the fuel mass flow $\dot{m}_B$ and its lower calorific value H_u:

$$\eta_{eff} = \frac{P_{eff}}{\dot{Q}_{add}} .$$

Diesel engines have at higher loads an effective efficiency of approx. 45%, large slow-running diesel engines have a much higher effective efficiency. Gasoline engines have, depending on the combustion process, an effective efficiency at their best point of approx. 42%.

Mechanical efficiency

The mechanical efficiency sets the effectively measured power P_{eff} in proportion to the pressure-indicated cycle power P_{ind}. The indicated power is determined from the work W – this is the area of the real pressure characteristic against volume $\int pdV$ – and the time t per working cycle according to the following relationship:

$$P_{ind} = \frac{dW}{dt} \approx \frac{\Delta W}{\Delta t} .$$

The effective power differs from the indicated power essentially in the frictional losses (piston, bearings), the power-transmission losses from control elements (camshaft, valves), and the power of the accessories (oil and water pumps, fuel-injection pump, alternator). For the mechanical efficiency:

$$\eta_m = \frac{P_{eff}}{P_{ind}} .$$

Normal mechanical efficiencies are load-dependent and at full load come in at just under 90%, while at low part load (10% load) values around 70% are registered.

Efficiency-of-cycle factor

The efficiency-of-cycle factor describes the efficiency with which the real cycle can be approximated by the chosen reference cycle. It therefore contains losses resulting in particular from dissipation losses. For a detailed loss analysis it is recommended to split the efficiency-of-cycle factor between the high-pressure loop and the charge-cycle loop (Table 2).

Normally, for calculation purposes, an ideal gas with temperature-dependent heat capacities is assumed and the cycle is used with geometrically identical dimensions, the same air ratio, without residual exhaust gas, complete combustion, and thermally isolated walls. An engine described in this way is also called a "perfect engine". The efficiency-of-cycle factor at full load is in the range of approximately 80 to 90%.

If an ideal gas with constant heat capacity is still to be used, it is possible to introduce the "efficiency of the perfect engine", which sets the power of the "perfect engine" in proportion to the power of the "ideal cycle".

Fuel-conversion factor

In particular, gasoline engines with rich-mixture combustion (air/fuel ratio $\lambda < 1$) are subject to high HC and CO emissions, which usually cannot be taken into account in the approach of added heat via the fuel calorific value H_{uB}. However, the exothermy of these gases H_u is considerable, which is usually also reflected in higher exhaust-gas temperatures after the oxidation-type catalytic converter. It is taken into account in the fuel-conversion factor:

$$\eta_B = \frac{(H_{uB} - H_u)}{H_{uB}} .$$

$\eta_B = 1$ is usually set for diesel engines. In gasoline engines the figure can drop to 0.95, and with very rich air ratios $\lambda < 1$ even further below that.

Efficiency chain

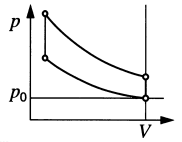
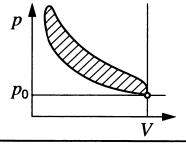
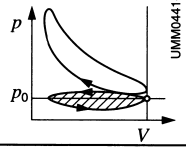
The entire efficiency chain can be described as follows (Table 2):

$$\eta_{eff} = \eta_i \, \eta_m = \eta_{th} \, \eta_g \, \eta_m .$$

References

[1] G. Woschni, Die Berechnung der Wandverluste und der thermischen Belastung der Bauteile von Dieselmotoren, MTZ 31 (1970).

Table 2. Graphic representations and definitions of the individual and overall efficiencies of the reciprocating-piston engine
The hatched areas indicate the newly added work component of the characteristic efficiency quantity. The efficiencies are explained in the text.

Pressure vs. volume diagram	Designation	Boundary conditions	Definition	Efficiencies
	Theoretical reference cycle, e.g. constant-volume cycle	Ideal gas, constant specific heat, infinitely rapid heat addition and dissipation, etc.	$\eta_{th} = 1 - \varepsilon^{1-\kappa}$ Theoretical or thermal efficiency	η_{th}
	Real high-pressure working cycle	Wall heat losses, real gas, finitely rapid heat addition and dissipation, variable specific heat	η_{gHD} Efficiency factor of the high-pressure cycle	η_g
	Real charge cycle (4-stroke)	Flow losses, heating of the mixture or the air, etc.	η_{gLW} Gas-exchange efficiency	
Mechanical losses, cannot be suitably represented in the pressure vs. volume diagram	Losses due to friction, cooling, accessories	Real engine	η_m	η_m

η_i

η_{eff}

Mixture formation, combustion, emissions

All internal-combustion engines share the common feature of combustion occurring after induction of the fresh mixture or air and after subsequent compression. In reciprocating-piston engines with internal combustion this takes place close to top dead center (TDC). The result is a pressure increase, which is transmitted via the piston and the connecting rod to the crankshaft in the form of a crankshaft torque (Figure 1).

The sequence of compression and subsequent combustion on the one hand has a significant impact on the pressure characteristic and thus on the efficiency and the torque output. On the other hand, this sequence defines the creation of emissions inside the engine. In this respect, gasoline and diesel engines differ in their process control.

Gasoline engine

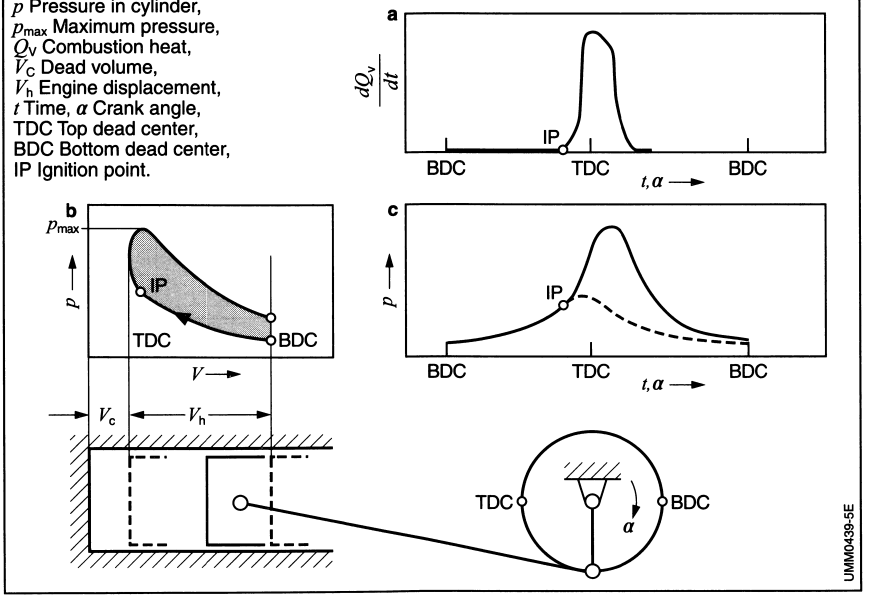
The characteristic feature of a gasoline engine is that it uses an external ignition source, normally an electrode spark plug. Ideally a suitable homogeneous air/fuel mixture is created to provide the required flammability. This is achieved by means of external mixture formation (intake-manifold injection) or internal mixture formation (gasoline direct injection).

Mixture formation

For the most part, a homogeneous mixture preparation is effected in a gasoline engine, i.e. the intake air is fully mixed with the vaporized or atomized fuel during the induction and compression strokes. The excellent vaporization qualities of gasoline enable it to be injected into the intake manifold. Modern stratified-charge combustion processes are, on the other hand,

Figure 1: The engine power cycle

a) Course of combustion,
b) Representation in the p-V diagram (enclosed area corresponds to indicated work),
c) Representation in the p-t or p-α diagram,
p Pressure in cylinder,
p_{max} Maximum pressure,
Q_V Combustion heat,
V_C Dead volume,
V_h Engine displacement,
t Time, α Crank angle,
TDC Top dead center,
BDC Bottom dead center,
IP Ignition point.



characterized by partially heterogeneous mixture formation (see Stratified-charge operation).

Mixture formation is decisively influenced by the vaporization conditions, the injection pressure, the cylinder-charge movement, and the time available with the aim of homogenization. Essentially, mixture formation involves the interaction of two processes: droplet vaporization caused by the temperature difference (Figure 2) and droplet disintegration by aerodynamic forces (Figure 3). Intake-manifold injection and direct injection differ here (Table 1).

Intake-manifold injection

In the case of intake-manifold injection, a film of mixture is created on the intake-manifold wall ahead of the intake valve whose fuel mass is reduced more intensively as the air velocity increases. This air velocity changes along broadly linear lines with the engine speed. On

Figure 2: Fuel vaporization

1 Diesel, 2 Petroleum, 3 Gasoline.

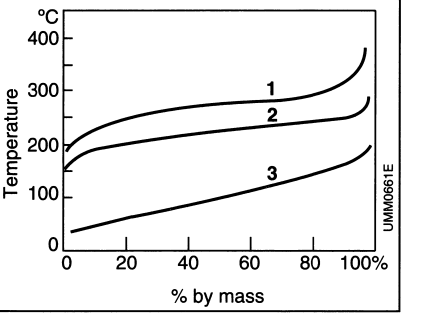
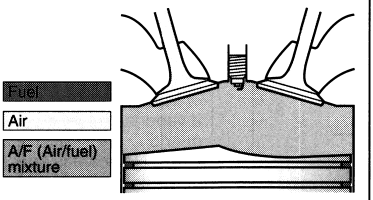
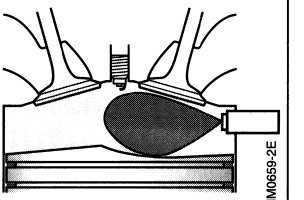


Table 1: Operating strategies of gasoline engines

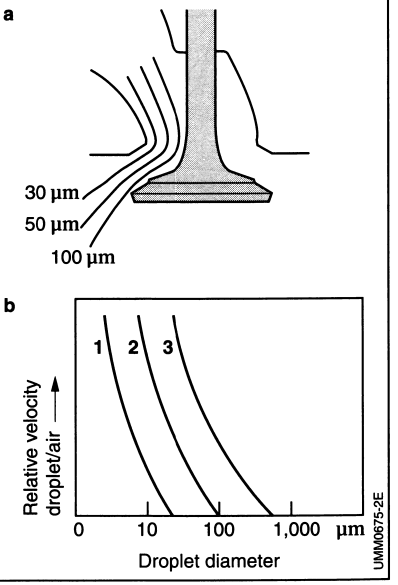
Operating strategy	Stoichiometric	Rich	Lean	Ultra-lean	Stratified-charge operation lean
Mixture formation	Homogeneous				Heterogeneous and homogeneous
Combustion-chamber composition					
Fuel injection	Intake-manifold injection and direct injection				Direct injection
Ignition	External ignition source	External ignition source	External ignition source	Auto-ignition	External ignition source
Typical compression ratio	8...12		11...13	12...16	11...14
Load control	Quantity				Quality
Operating range	Entire program map	Full load, high speed range	Entire program map	Part load	Part load
Application, development stage	Conventional, series		Gas engines, series	Research stage	New combustion processes

account of the low temperature and incomplete vaporization in the intake manifold with the resulting formation of film, intake-manifold injection occurs at a very low injection-pressure level below 10 bar.

The dynamics of the wall-applied fuel film and the mechanisms involved during vaporization are one of the main causes of inaccurate fuel metering, above all during transient engine operation. Only the smaller droplets entrained with the intake flow reach the inside of the cylinder (Figure 3). Typically, their characteristic diameter is already less than 30 µm. Here, droplet acceleration is proportional to the ratio of relative velocity to the air and to the droplet diameter.

The very high turbulence intensity and the high flow velocities in the valve gap result in very good mixture formation.

Figure 3: Droplet formation
a) Characteristic flow situation in valve gap,
b) Interaction between droplet diameter and relative velocity of droplets with regard to air flow.
1 0 %,
2 50 %,
3 Over 90 %.



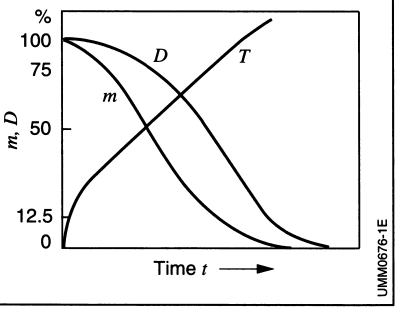
As process control progresses the remaining small fuel droplets assume the temperature of the mixture and vaporize (Figure 4). This enables the vaporized fuel to distribute more effectively in the combustion chamber, resulting in homogenization. An optimum combustion-chamber design prevents intensive contact of fuel with the wall, since this always involves the risk of fuel condensation.

Gasoline direct injection

Gasoline direct injection does not utilize the mechanisms of mixture formation in the valve gap. A higher injection pressure of 50 to 350 bar is therefore required. Injection is completed by no later than the gas-flow bottom dead center – i.e. at the end of the induction stroke – to allow sufficient time for homogenization during the compression stroke. Multiple injections are also carried out in the induction and compression strokes in order to achieve better homogenization. This increases the charge movement, which in turn encourages mixing. In heterogeneous combustion processes injection occurs only at the end of the compression stroke (see Stratified-charge operation).

The mixture included in the combustion chamber is now compressed, depending first and foremost on the throttle-valve position and the compression ratio, to

Figure 4: Droplet vaporization
m Mass, D Diameter, and T Temperature of a vaporizing fuel droplet.
m and D are shown scaled to the maximum value. No scale is specified for T, because only the basic characteristic is involved.



a pressure level of 10 to 40 bar. This equates to a temperature level of 300 to 500 °C, depending first and foremost on the ambient temperature and the compression ratio.

The advantage of direct injection is that the fuel is metered exactly. Vaporization of the fuel in the combustion chamber also requires the cylinder charge to be adequately cooled. This enables the compression ratio to be increased by roughly one unit, resulting in greater efficiency.

Combustion in gasoline engines

In all combustion processes combustion – oxidation – occurs only at the end of the compression phase and in the early expansion phase. The subsequent combustion process differs depending on the nature of mixture formation (homogeneous or heterogeneous). An entirely homogeneous mixture reacts in a premixed combustion, an entirely heterogeneous mixture in a mixture-controlled combustion. During the stratified-charge operation of modern direct-injection engines, a large proportion (>50 %) of the injected fuel is likewise homogenized up to the start of combustion.

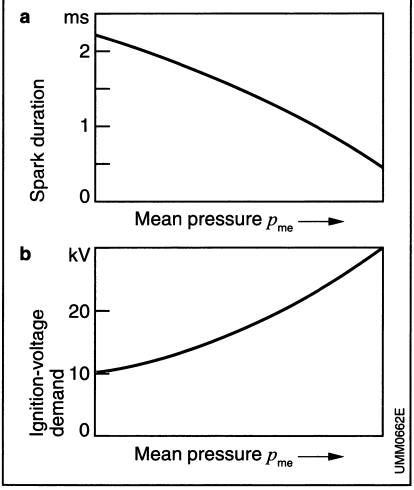
In both homogeneous and partly heterogeneous mixture formation, actual combustion is preceded by ignition and the ignition phase.

Ignition

Ignition is typically performed with the aid of an electrode spark plug. When a high voltage is applied, a sparkover occurs between the electrodes, depending on the mixture state (i.e. pressure, temperature, and mixture composition), the electrode gaps, and the air humidity. The high voltage is typically in the range of 10 to 40 kV. First and foremost the number of molecules between the electrodes influences the ignition-voltage demand (Paschen's law). The mixture ignited by the spark must in the course of its combustion release that amount of energy that is needed to ignite the immediately adjacent mixture. As leaning of the mixture increases, the energy content of this mixture decreases when the electrode gap remains constant. This is accompanied by an increase in the energy demand needed to ignite the adjacent – also lean – mixture. By enlarging the electrode gap, it is possible to increase the volume ignited by the spark and thereby raise the energy content. However, enlarging the electrode gap requires an increase in the ignition voltage. In this way, this increases for example in lean-combustion processes or in the event of a load increase. In the event of a load increase, the spark duration simultaneously decreases as the ignition-voltage demand increases (Figure 5). On account of heat losses at the spark-plug electrodes, heat-convection losses and cyclically fluctuating mixture states, the electrical ignition energy is above the theoretical minimum ignition energy by up to an order of magnitude (Figure 6). The stochastically fluctuating states (flow field and mixture state) between the electrodes are the main cause of large cyclic variations in a gasoline engine. This situation is improved by enlarging the electrode gap. Modern engines here already operate with a maximum value above 1 mm. Enlarging the electrode gap requires an increases in the ignition voltage.

Figure 5: Ignition sparks and ignition-voltage demand

a) Spark duration,
b) Ignition-voltage demand U_Z.
The ignition-voltage demand rises as exhaust-gas recirculation or leaning increases.



The purpose of ignition is to ignite the air/fuel mixture and thereby initiate the actual combustion process. Depending on the velocity of the subsequent combustion and on the piston speed (and thus on the engine speed), the point of ignition must be variably adapted (Figure 7).

Figure 6: Minimum ignition energy for a methane/propane mixture

1 Ignition energy for a static mixture,
2 Ignition energy for a mixture with a flow velocity of 6 m/s,
3 Ignition energy for a mixture with a flow velocity of 15 m/s.
With an increase in the proportion of inert gas the ignition-voltage demand increases for all the flow states shown.

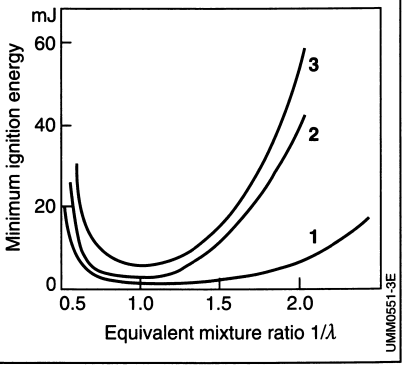
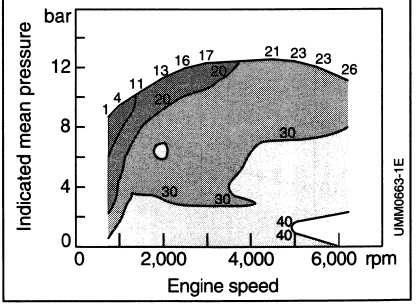


Figure 7: Program map of ignition point

Values in °CA before TDC for a homogeneously operated gasoline engine.
Indicated mean pressure is obtained from the indicated work/(engine speed times engine displacement).



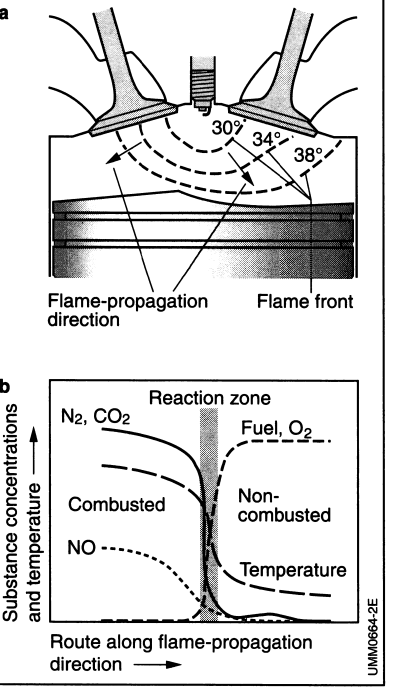
Combustion of homogeneous mixtures

Propagation of the flame front

Ignition sets in progress the actual combustion in a homogeneously operated gasoline engine. Here the flame propagates from the spark plug. It is possible to define a flame-front velocity which is composed of the sum of the flame velocity and the flame-front movement (charge movement, expansion caused by density differences). In a homogeneously operated gasoline engine a good distinction can be made between the combusted and non-combusted mixture, since the flame continuously propagates outwards (Figure 8).

Figure 8: Flame propagation and premixed combustion

a) Flame-front propagation in a gasoline engine during homogeneous operation (the figures in degrees are calculated from the ignition point),
b) Temperature and substance concentrations along the flame-propagation direction.



It takes a few milliseconds for the flame to reach a radius of around one centimeter in order to propagate unhindered from the piston recess and cylinder head at a velocity well in excess of 10 m/s.

Influences on flame velocity
The decisive factor now is the propagation velocity of the flame, which is also referred to as the turbulent flame velocity. The faster the flame propagates, the better the combustion inside the engine. The following factors encourage a high flame velocity: a low proportion of inert gas, a temperature increase in the non-combusted mixture, a pressure increase, and a high turbulence level.

Most fuels have their maximum flame velocity in slightly rich-burn operation in the range of λ = 0.85 to 0.9. A further advantage of slightly rich-burn engine operation is the cooling effect provided by the excess fuel. Racing engines and passenger-car engines in the rated-output range are therefore operated with slightly rich mixtures.

Increasing the proportion of inert gas reduces the flame velocity. A practical application of inert-gas variation is exhaust-gas recirculation (EGR), in which combusted exhaust gas with the chief constituents CO_2, H_2O and N_2 is added to the air/fuel mixture. As a rule of thumb an exhaust-gas recirculation rate of 10 % already reduces the flame velocity by 20 %.

The reason why modern gasoline engines can nevertheless be operated with high internal exhaust-gas recirculation rates is the influence of temperature. Doubling the temperature increases the flame velocity by a factor of roughly 4.

Cylinder pressure has a lesser influence; an increase in pressure causes a slight acceleration of the flame velocity.

Turbulences in the combustion chamber
The turbulence level in the combustion chamber has the greatest influence on the combustion velocity. The flame velocity varies along roughly linear lines with the turbulence intensity. Turbulence intensity is a measure of the high-frequency change in flow velocity at a given point in the combustion chamber [1]. The turbulent kinetic energy is proportional to the square of the turbulence intensity.

Turbulence intensity is a three-dimensional quantity which is influenced above all by the flow profile of the charge movement in the combustion chamber. The velocity of the flow processes inside the engine that increases linearly to the greatest possible extent as engine speed increases is extremely important. As the flow velocity increases, so too does the turbulence intensity in the combustion chamber. This is the reason for stable engine operation over a very wide engine-speed range. The flame would otherwise no longer be able to burn cleanly at increasing engine speed and consistent flame velocity because of the shorter time available for combustion. However, the positive contribution of turbulence cannot equal completely the influence of speed, which is why combustion at high speeds nevertheless stretches over a wider crankshaft range. This is an additional reason why gasoline engines operate with less efficiency at higher speeds.

In gasoline engines the turbulence in the combustion chamber is vitally important to energy conversion. The decisive factor in causing turbulence is the charge movement in the cylinder, which is fundamentally influenced by the intake flow (depending on the configuration of the ports in the cylinder head) and the combustion-chamber shape (Figure 9).

Combustion causes a pressure increase which can also be simultaneously heard. For comfort reasons, suitable combustion-adjustment measures must be taken to minimize this pressure increase as much as possible. This poses a conflict of objectives with thermodynamic efficiency. The maximum pressure-increase gradient of gasoline engines is in the range of between 0.5 and 3 bar/°CA.

Combustion of partially homogeneous mixtures

Modern stratified-charge combustion processes facilitate operation with excess air in the part-load range. Effective mean pressures of $p_{me} < 1$ bar make even operation with a mean excess-air factor $\lambda > 5$ possible in stratified-charge operation. The main benefit here is the improved charge cycle, because throttling (which has a negative impact on overall efficiency) can be largely dispensed with.

Conventional operation is not possible without throttling in the lower part-load range, because homogeneous mixtures leaned in this way burn much too slowly and therefore incompletely. The solution is to provide local fuel stratification in the area of the spark plug by means of an optimized injection strategy in the late compression phase. The difficulty here lies in optimally matching the injection strategy and ignition, since the conditions between the electrodes change (Figure 10).

Figure 9: Turbulent and mean kinetic energy as a function of the crankshaft position, each referred to mass

1 Turbulent kinetic energy,
2 Mean kinetic energy.

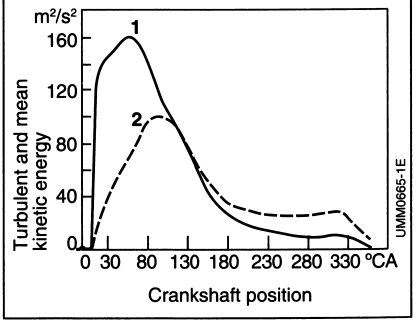


Figure 10: Excess-air factor λ over time in a spherical volume (with radius r) around the electrode midpoint

1 Stratified-charge combustion process, λ within $r = 2$ mm (gaseous fuel),
2 Stratified-charge combustion process, λ within $r = 3$ mm (gaseous fuel),
3 Stratified-charge combustion process, λ within $r = 5$ mm (gaseous fuel),
4 Stratified-charge combustion process, λ within $r = 5$ mm (liquid and gaseous fuel),
5 Optimized stratified-charge combustion process,
6 Homogeneous combustion process.

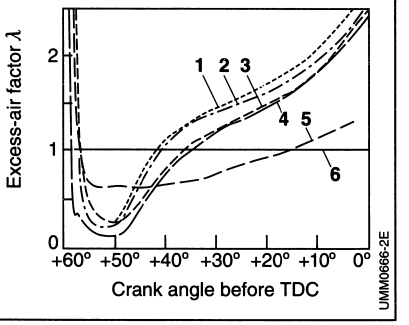
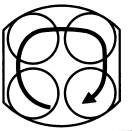
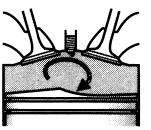
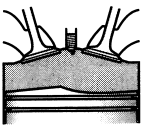
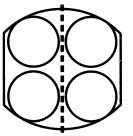
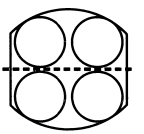


Table 2: Charge cycle and flow profile

Operating strategy	Swirl	Tumble	
Flow			
Location of flow axis	 Vertical axis z	 Lateral axis x	 Lateral axis y

UMM0660-1E

Charge movement
Large-scale, swirling and circular flows with diameters similar to the characteristic quantities of the combustion chamber provide the basis for the cylinder-charge movement. A basic distinction is made between flows about the vertical axis (cylinder axis), which are called swirl, and flows about the two lateral axes (crankshaft axis and perpendicular to this), which are called tumble (Table 2). In reality there is an overlap of the three basic flows, which gives rise to complex three-dimensional flow fields. Tumble and swirl essentially differ in their behavior inside the engine.

Tumble flow disintegrates up to the moment when compression dead center is reached and contributes primarily to flame propagation in the first combustion half.

Swirl flow lasts longer into the subsequent expansion phase. The disintegration of the large swirls into ever-decreasing, turbulent structures encourages the creation of turbulence. In the continuing process, however, the viscosity of the working medium results in a disintegration with a detrimental effect on the combustion velocity.

The creation of turbulence is assisted by the combustion-chamber geometry. In particular, flows in the piston-recess or squish-gap area (area between piston and cylinder head) support flame propagation (see e.g. [2]).

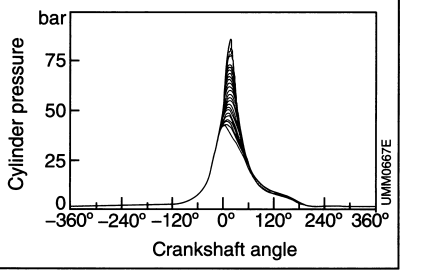
One of the greatest physical challenges associated with fashioning homogeneous combustion processes is the combustion behavior in the expansion phase, since typically over 10 % of the supplied fuel is still not converted 30° CA after TDC. At this point the non-combusted mixture is still only in the immediate vicinity of the wall and must still be converted after re-emerging from the top-land area between piston ring and cylinder wall (see e.g. [2]). This phenomenon results in incomplete combustion in the final phase.

Uncontrolled combustion
Numerous undesirable processes make it difficult to fashion a homogeneous combustion process. In addition to cyclic variations, knock and auto-ignition characteristics have detrimental effects. Extreme forms of advanced ignitions can occur, particularly in modern, highly turbocharged engines.

Cyclic variations
Cyclic variations are a consequence of the premixed combustion of homogeneous combustion processes which reacts sensitively to numerous disturbance values (Figure 11). Such disturbance values are, for example, the mixture composition, the residual-exhaust-gas content, the thermodynamic state variables, and the flow profile. All the values vary from working cycle to working cycle and cause noticeable cyclic fluctuations in energy conversion. In particular, leaning with excess air of more than 40 % increases the cyclic variations significantly. Initiation of ignition has the greatest influence on cyclic variations.

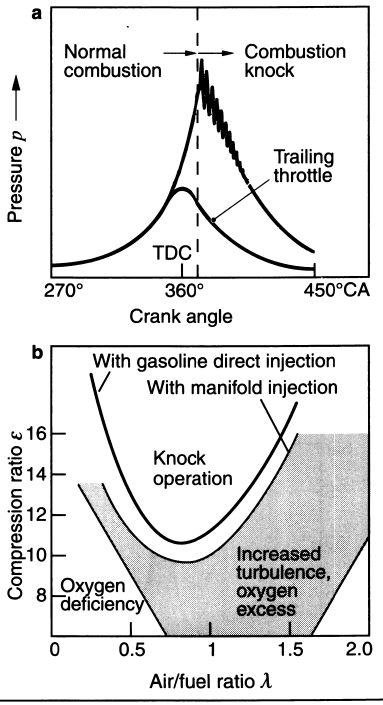
Knocking
Another challenge is posed by the undesirable phenomenon of knocking. Knocking occurs in response to increases in pressure and temperature after the onset of regular combustion until the local conditions stimulate auto-ignition. Here, the residual mixture in response to the temperature increase as a result of the

Figure 11: Cyclic fluctuations of the cylinder pressure of the individual working cycles during stoichiometric operation ($\lambda = 1$)



progressing combustion reaches the ignition temperature and burns almost simultaneously without a further controlled flame propagation. The pressure pulsations generated in the process result in material wear in the engine bearings and cause serious engine damage in the event of prolonged operation. The temperature peaks can also result in damage to components. Typically, knocking only then occurs in a working cycle if 80 % of the fuel in the combustion chamber has not been combusted. Knocking can be observed particularly at low engine speeds with suitably sufficient time for auto-ignition and at high loads with high combustion-chamber temperatures. Fuels with high ignition temperatures such as, for example, methane or ethane reduce knock sensitivity. Knocking combustion can be reduced by shifting the energy conversion in the retardation direction (ignition timing towards TDC). High-compression and turbocharged engines are more knock-sensitive because of their higher final compression temperatures (Figure 12). Useful measures to counter knocking include effective cylinder cooling of the hot areas, for example also by the vaporization effect of gasoline direct injection, increasing the turbulence, reducing the geometrical compression ratio, and optimizing the fuel, for example with additives.

Figure 12: Compression ratio, knocking and knocking combustion
a) Example of knocking combustion,
b) Effect of compression ratio and air/fuel ratio on the operating limits (knocking and engine misfires). Direct injection instead of intake-manifold injection and increased turbulence shift the operating limits.



Auto-ignitions
In contrast to knocking, auto-ignitions can occur quite perfectly even in very retarded compression situations. Possible causes of auto-ignitions are, for example:
- Over-retarded moment of ignition with incomplete combustion, thus fuel-film formation causing auto-ignitions,
- Full-load operation with high cylinder-component temperatures,
- Ignition caused by abrasion and hot particulates,
- Oil emissions caused by defective or worn piston rings.

Extreme forms of auto-ignition can occur in turbocharged and high-compression gasoline engines; these can result in peak pressures in excess of 150 bar and serious damage. However, these extreme auto-ignitions occur only very rarely with a probability of less than 0.01 per thousand.

Pollutant formation and pollutant reduction in gasoline engines

In addition to the unavoidable combustion products of carbon dioxide (CO_2) and water (H_2O), the concentration of which is dependent on the composition of the fuel, nitrogen oxides (NO_x), non-combusted hydrocarbons (HC) and carbon monoxide (CO) are the main emissions from gasoline engines (Figure 13).

Pollutants

Nitrogen oxides

Nitrogen oxides (NO_x) require four factors in order to be created: oxygen, nitrogen, high temperatures, and time. Because oxygen and nitrogen are determined by the mixture composition in gasoline engines and the available time is defined by way of the engine speed, nitrogen oxides in gasoline engines can only be reduced by low maximum combustion-chamber temperatures (e.g. by means of ignition retardation and exhaust-gas recirculation).

Hydrocarbons and carbon monoxide

Hydrocarbons (HC) and carbon-monoxide emissions (CO) are the result of incomplete combustion. In rich-burn operation there is, on account of the lack of oxygen, an increase in both HC and CO emissions. In lean-burn operation with accordingly reduced flame temperature a more intensive extinguishing of the flame is encountered, above all in the fire-land area close to the wall, with a resulting increase in HC emissions. Because of the excess oxygen the slightly oxidizable CO emissions are nevertheless reduced.

Soot and particulates

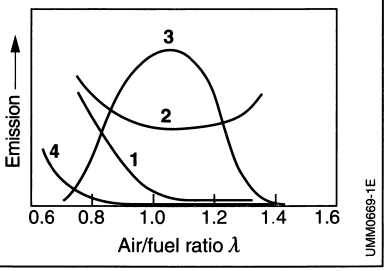
Engines with intake-manifold injection cause – primarily in extreme rich mode – increased soot emissions. These are however lower compared with gasoline engines with direct injection. With the Euro 6c emission-control standard the particulate limit for gasoline engines with Direct Injection has been reduced to 1/10 of the value applicable to Euro 6b such that a particulate filter must be used in many vehicle models.

Sulfur oxide

Sulfur-oxide emissions are dependent on the sulfur content in the fuel. Currently only a sulfur content of 10 mg/kg is permitted in Europe so that sulfur-oxide emissions are of secondary importance.

Thanks to exhaust-gas treatment, modern homogeneous gasoline engines are, once they have reached the catalytic-converter operating temperature, virtually zero-emission machines. Three-way catalytic converters in operation at $\lambda = 1$ reduce the nitrogen-oxide emissions while simultaneously oxidizing HC and CO molecules (see Three-way catalytic converter). Alternative approaches are required in lean-burn operation. For this reason, NO_x storage catalytic converters are typically used for stratified-charge concepts. These catalysts store the nitrogen oxides. Rich-burn engine operation applied at regular intervals reduces the stored nitrogen oxides at high temperatures. Because NO_x storage catalytic converters are sensitive to sulfur contamination, desulfation cycles must additionally be completed in slight rich-burn operation at temperatures in excess of 600 °C (see NO_x storage catalytic converter).

Figure 13: Gasoline-engine emissions
1 Carbon monoxide (CO),
2 Hydrocarbons (HC),
3 Nitrogen oxides (NO_x),
4 Soot.



Load control in gasoline engines

In homogeneously operated gasoline engines the load is adjusted by means of the injected fuel mass. The corresponding air mass is adapted by means of the throttle-valve position for the required operation at $\lambda = 1$. This is known as quantity control. In the part-load range this causes induction throttling, which is detrimental to the overall efficiency. The cause of this is the work done during induction and emission of the gas, the share of which must be subtracted from the energy generated in the combustion process. This disadvantage can be partially compensated for by varying the valve timing. Advanced or retarded "intake closes", reduced valve lift or retarded "exhaust closes", in which hot exhaust gas is additionally inducted, are typical measures. Alternatively, external exhaust gas can be recirculated (exhaust-gas recirculation) in order to reduce throttling. In turbocharged gasoline engines the air flow rate and thus the fuel-mass flow rate is adjusted in the upper load range typically by means of the position of the wastegates on the turbocharger (see Exhaust-gas turbochargers).

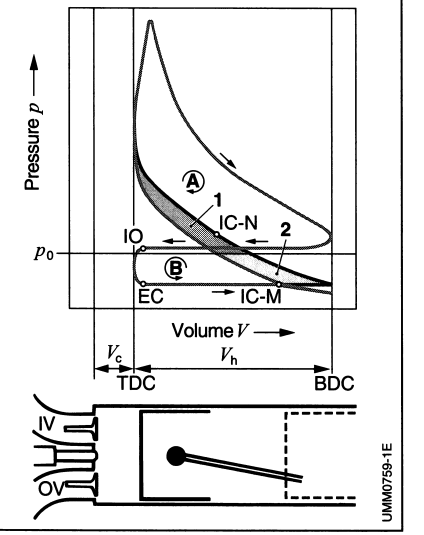
In engines with stratified charge the injected fuel mass is adjusted by the load in unthrottled operation at part load. This is known as quality control. The transition between stratified- and homogeneous-charge operation in the mid-load range requires control-engineering complexity and expenditure.

Power yield and efficiency

The part-load performance of a gasoline engine deteriorates as a result of charge-cycle losses (throttling), poor process control (peak pressures below 30 bar), and an increasing amount of engine friction. Because even at passenger-car driving speeds in excess of 100 km/h most vehicle engines still operate in the part-load range, measures to increase the efficiency in this operating range prove highly successful. Examples of such measures are:

Figure 14: Miller process: Power yield and efficiency increase by dethrottling via valve timing

1 Additional useful work,
2 Less charge-cycle losses.
TDC Top dead center,
BDC Bottom dead center,
IV Intake valve,
EV Exhaust valve,
IO Intake opens,
IC-N Intake closes (normal operation),
IC-M Intake closes (Miller process),
EC Exhaust closes,
V_c Compression volume,
V_h Cylinder displacement,
p_0 Atmospheric pressure (1013 mbar),
A High-pressure loop,
B Charge-cycle loop.



- Displacement reduction (downsizing),
- Cylinder shutoff (e.g. in V8 and V12 engines),
- Dethrottling (stratified charge, exhaust-gas recirculation, valve timing),
- Increased compression ratio,
- Longer transmission ratio to lower the engine-speed level.

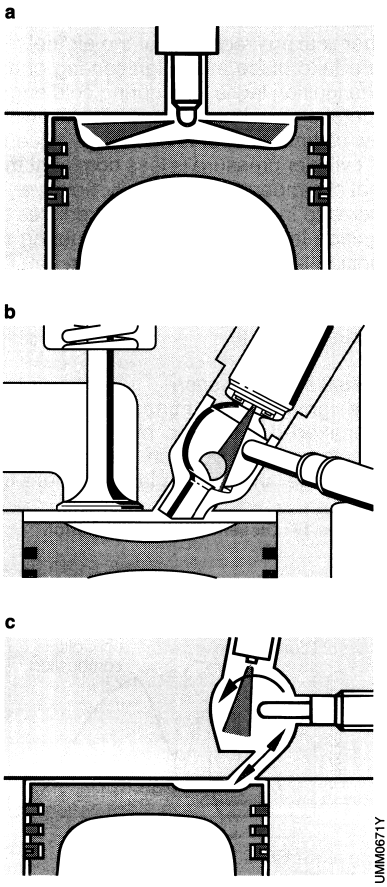
In the Miller process the intake valve is closed at a very advanced stage. In this way, the throttle valve can remain open and the throttling losses are lower (Figure 14).

Diesel engine

A significant characteristic feature of diesel engines is the absence of external ignition. This is achieved by injecting the ignitable fuels into highly compressed and therefore hot air. High final compression temperatures and pressures to levels in excess of 600 °C and 100 bar in turbocharged engines provide for extremely stable engine operation. Fuel-spray formation, fuel vaporization, fuel mixing, and subsequent combustion can take place within a very short period of time.

Combustion processes with direct injection have in the last few decades established themselves as the favored choice over indirect combustion processes such as whirl-chamber or prechamber systems. In indirect combustion processes the local flow in the secondary volume is essentially responsible for fuel preparation (Figure 15).

Figure 15: Diesel-engine combustion processes
Combustion-chamber shape and nozzle arrangement for:
a) Direct injection,
b) Prechamber system,
c) Whirl-chamber system.



Mixture formation

Mixture formation is dominated by the interaction of the injection spray with the flow field in the combustion chamber. Here, the challenge consists in quickly injecting and preparing relative large masses of fuel of up to 200 mg per liter displacement. A typical injection duration is in the range of 1 ms. The term used in relation to the fuel mass flow into the combustion chamber is the injection rate (unit: kg/s). This fuel mass flow is typically injected by multihole injection nozzles.

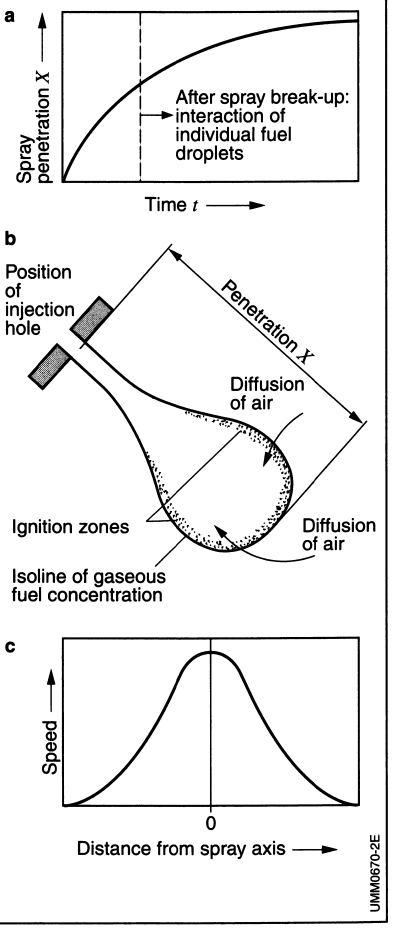
A combination of between four and ten injection holes with a diameter of between 120 and 150 µm is commonly used. As well as the small hole diameter, the high injection pressure of over 2000 bar encourages fast fuel injection and mixture formation.

The diameter of a characteristic spray is initially the same as that of the nozzle hole. After it has traveled a few millimeters, the spray breaks up into individual droplets which interact with the flow field. The liquid phase of the fuel spray can, especially depending on the density of the working medium (air, air-fuel mixture or air with EGR), penetrate a few centimeters into the combustion chamber before the spray is completely atomized or vaporized (Figure 16).

Turbulences encourage the formation of droplets and the vaporization of the fuel. In modern diesel engines more than 80 % of the turbulence intensity in the area of the fuel spray is generated by the injection. This is further assisted by the charge movement, whereby diesel engines with flat cylinder head experience predominantly more swirl than tumble

flows. An additional contribution can be made by compression-induced air flows from the outer to the inner area of the combustion chamber ("squish flows") or by the combustion-chamber design, e.g. by vaporization-inducing contact with the hot piston-recess area.

Figure 16: Spray propagation and mixture formation in diesel engines
a) Characteristic spray penetration behavior,
b) Typical injection-spray formation (the lean edge areas with infinite λ are outside the continuous line),
c) Velocity profile of injection spray.



Combustion in diesel engines

Combustion in diesel engines differs from conventional gasoline-engine operation in the compression-ignition behavior. In all, diesel-engine combustion can always be described in three consecutive processes: ignition lag, premixed combustion, and mixture-controlled combustion. Depending on the operating state and the map range, the three processes have different time components (Figure 17).

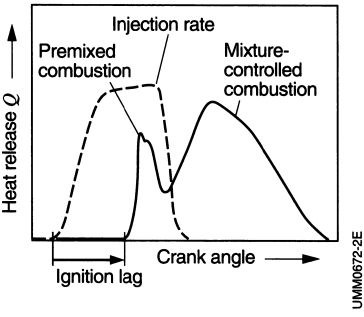
Ignition lag

Ignition lag refers to the period of time between start of injection and start of combustion. It is crucially dependent on the cylinder temperature, the cylinder pressure, and the combustibility of the fuel. Both mixture formation and the first chemical pre-reactions of the air/fuel mixture take place in the ignition-lag phase. The ignition lag is long during cold engine operation or when a low-quality fuel with a low cetane number is used. The influence of cylinder pressure is less dominant than that of temperature. Basically, however, an increase in cylinder pressure reduces the ignition lag. The fuel injected during the ignition-lag phase does not burn yet. The ignition lag can be between 0.1 ms in the rated-output range to over 10 ms in the cold-start range.

Premixed combustion

The ignition lag defines the phase of premixed combustion by means of the injected and as yet non-combusted fuel. The longer the ignition lag, the more fuel

Figure 17: Diesel-engine combustion



is premixed in combusted form. This fuel mass can perfectly well be over 20 mg per liter displacement. Combustion typically starts at the edge of the fuel spray, where the fuel has mixed very well with air and therefore optimum ignition conditions are present in terms of temperature and excess-air factor λ. The exothermic reaction results in a local temperature increase with temperatures above 2300 K, which quickly converts the as yet non-combusted premixed fuel in a chain reaction. Here, the chemical reactions that take place determine the rate of combustion. The self-accelerating chain reaction results in an extremely fast combustion with large pressure-increase gradients. For this reason, the premixed converted fuel mass must be kept to a minimum in diesel engines. This is typically achieved by means of a pilot injection whose local combustion exerts an initial temperature increase with a reducing effect on the ignition lag of the subsequent main injection.

Mixture-controlled combustion

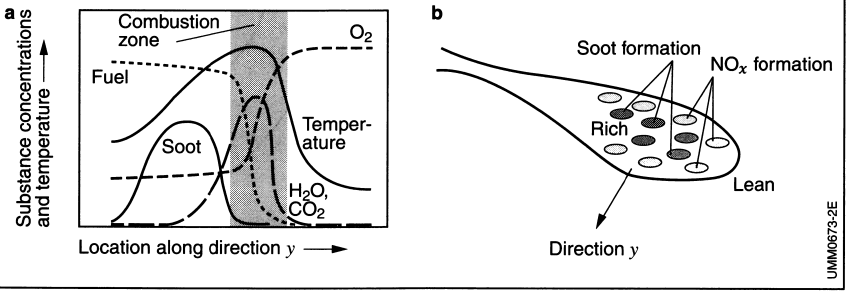
The amount of premixed converted fuel can be between less than 1 % in the upper full-load range and 100 % in the lowest part-load range. The remaining amount burns in a mixture-controlled combustion. In contrast to premixed combustion, in the case of mixture-controlled combustion – also called diffusion combustion – the transfer of oxygen into the combustion zone defines the conversion rate. It is difficult to distinguish between a combusted zone and a non-combusted zone, because there is no precisely defined flame front. Basically, the diffusion flame establishes itself at the edge of the spray and burns in a limited range at $0.8 < \lambda < 1.4$. As the boundary conditions change (e.g. further fuel vaporization, oxygen transfer, wall contact), so too the reaction zone strays to where locally stoichiometric conditions prevail (Figure 18).

Turbulences

Mixture-controlled combustion dominates in the high load range, in which large quantities of fuel are injected. Here, the mixture-formation and combustion processes occur in parallel. As with premixed combustion, the conversion rate can also be influenced by the injection. A lesser influence, but also of an accelerating character, is exerted by increases in both temperature and pressure and by a reduction in the proportion of inert gas. The dominant factors are mixture formation and the transfer of oxygen into the combustion zone by a high local turbulence.

For this reason, turbulence intensity is the deciding variable in fashioning diesel-engine combustion processes and is reflected in ever higher injection pressures with a high kinetic energy on the part of the fuel spray, which is subsequently converted into turbulent kinetic energy. Local turbulence causes the vital transfer of oxygen into the local reaction zones. The cylinder-charge movement (swirl, squish flows) supports this phenomenon, the main contribution however being made by the pulse of the injection spray. As well

Figure 18: Mixture-controlled combustion
a) Substance concentrations and temperature distribution,
b) Formation of pollutants inside the injected spray.



as increasing the injection pressure, increasing the hole diameter or the number of holes is also conceivable. However, an increase in the injection rate results for the most part in a local over-enrichment with a negative impact on fuel conversion.

Particular characteristics of diesel-engine combustion

Cold-start processes

Cold-start processes in diesel engines pose a particular challenge, particularly at outside temperatures below –10 °C. At starter speeds down to below 100 rpm a large proportion of the cylinder charge escapes through the piston rings during the relatively slow compression phase. In addition, the cold cylinder temperatures give rise to increased wall heat losses. Results experienced are low peak pressures below 30 bar and, depending on the outside temperature, low peak temperatures below 400 °C.

Fuel vaporization at top dead center

Fuel vaporization at top dead center causes further cooling. The outcome is very long ignition lags. In an extreme-case scenario there may be no ignition whatsoever with the result that fuel can accumulate in the cylinder on the cylinder wall and piston over several working cycles. Its ignition after a few working cycles can, on account of the fuel mass accumulated in the meantime, result in very high peak pressures above 150 bar.

Friction losses

Because the cold-start phase does not provide for enough time for an adequate hydrodynamic lubricating film to be built up at the appropriate bearing points of the crankshaft drive, the result is a negative effect on the mechanical engine systems. Typical measures to help improve the cold-start process include therefore preheating the intake air, the oil or the water. The latter, as well as influencing the combustion-chamber temperature, above all brings about lower engine friction and increased starter speeds.

Height dependence

A further phenomenon is encountered during operation at very hot outside temperatures or at high altitudes above 1000 m. Because the air has a lower density, its mass in the cylinder is reduced. The effect on combustion is initially not decisive. However, the exhaust-gas temperature rises as a result of the lower amount of excess air.

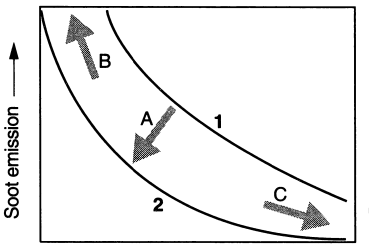
This phenomenon is also encountered with turbocharged engines. A power reduction above all during operation at altitude can therefore be a necessary application measure.

Deposits in the injection nozzles

A power drop in the order of 1 to 3 % is often observed in diesel engines after a break-in period. This is caused by the fuel-injection system. Deposits in the injection nozzles cause a slightly reduced nozzle-hole diameter, resulting in a smaller mass flow and thus a power loss. These deposits may be caused, for example, by copper, zinc or oil contaminants in the diesel fuel.

Figure 19: NO_x and soot emissions

1 Upper limit curve,
2 Lower limit curve.
A Decrease in soot and NO_x as a result of: exhaust-gas recirculation and injection-pressure increase, partial homogenization, H_2O injection.
B Increase in soot and decrease in NO_x as a result of: ignition retard, injection-pressure reduction, reduction of O_2 concentration, Miller process.
C Decrease in soot and increase in NO_x as a result of: ignition advance, injection-pressure increase, increase in O_2 concentration.



Pollutant formation and pollutant reduction

In contrast to gasoline engines, which are characterized with the introduction of the three-way catalytic converter during operation at $\lambda = 1$ by extremely low emissions, pollutant reduction inside the engine plays a much more important role in diesel engines. In addition to the emissions familiar from gasoline engines, i.e. CO_2, H_2O, NO_x, HC and CO, soot and particulate emissions must also be taken into consideration.

Pollutants

Nitrogen oxides

Measures which lower the combustion temperatures are useful in reducing the nitrogen oxides. This can be done effectively by reducing the oxygen concentration in the combustion zone. The combustion temperatures can also be reduced very easily by retarding injection or reducing the injection pressure.

Temperature-reducing measures such as exhaust-gas recirculation, Miller processes (see e.g. [2]) or partial homogenization reduce the nitrogen-oxide levels. This overcompensates for the slight increase in soot emissions that are often observed as well (Figure 19). The level of complexity and expenditure needed to reduce both these components is high. The measure currently used involves increasingly higher exhaust-gas recirculation rates for nitrogen-oxide reduction in combination with very high injection pressures (> 2000 bar) for soot reduction.

Soot and particulates

Reducing the injection pressure or the oxygen concentration typically causes an increase in soot emissions. The formation of soot is a complex process and is dependent on both fluid-dynamic and thermodynamic boundary conditions. Initially, the locally highly rich zones ($\lambda << 1$) produce considerable amounts of soot, which are then however substantially reduced by over 70 % in the subsequent combustion by oxidation processes. A high turbulence level, which encourages the soot to oxidize in the expansion phase, is crucially important. But the temperature level too is important. On account of important local interactions between the injection spray, the combustion zone, the non-combusted mixture and the recess geometry, how the combustion process is fashioned has a decisive impact on the production of emissions.

A distinction is made between soot and particulate emissions in this context. Soot consists of pure carbon compounds, whereas particulates also include fuel or lube-oil droplets, ash, rubbed-off metal, corrosion products, and sulfate compounds.

HC and CO compounds

HC and CO compounds are typically uncritical in diesel-engine emissions. The effect of hydrocarbons on particulate emissions must be taken into consideration. HC and CO increase in concentration in particular during a highly retarded combustion with incomplete burn behavior.

Mixed forms and alternative operating strategies

The classic diesel-engine operating strategy is characterized by one or more injections in the TDC range. Established gasoline-engine combustion processes are characterized by homogeneous or partially homogeneous (with charge stratification) operation. However, alternative forms of process control, which sometimes cannot be clearly assigned to the diesel-engine or the gasoline-engine process, are currently being developed.

Homogeneous-charge compression ignition in diesel operation
With the HCCI combustion process (Homogeneous Charge Compression Ignition), which is published in a wide range of variations [3], the motivation is, through the application of very advanced injection (at least 40 to 50° CA before TDC), to achieve homogenization, significant leaning and therefore a NO_x reduction. Reliable ignition nevertheless takes place on account of the high diesel-engine compression temperature. The compression ratio must be lowered to 14 to 16 in order to maintain control of the combustion process. Exhaust-gas recirculation is typically used to enable the temperature level in the cylinder to be increased at low loads. Nevertheless, obtaining this over the entire program map, particularly in the full-load range, is very demanding, since the pressure-increase gradients become extremely high and control of transient engine operation is highly complex in view of all the engine states that arise.

Compression ignition in gasoline operation
Gasoline-engine combustion processes have been further developed along similar lines to diesel-engine HCCI operation in order to achieve in the part-load range dethrottled lean-burn operation with corresponding fuel-consumption advantages over conventional, stoichiometrically operated engines. The disadvantages of lean-burn operation with regard to conversion in the catalytic converter are compensated by extremely low untreated NO_x emissions caused by leaning. Reliable ignition of the in itself hardly combustible mixture is achieved by a high compression ratio above 13. Optimally the compression ratio is variable and can be further decreased up to full load on account of the increased combustion-chamber temperatures.

Gasoline-engine stratified charge
Direct-injection, stratified-charge gasoline-engine combustion processes share much in common in terms of combustion with classic diesel-engine operation and therefore represent a mixed form of conventional gasoline-engine and diesel-engine combustion. Combustion processes of this type are becoming increasingly established thanks to their efficiency advantages by dethrottling in part-load operation.

Multifuel engines
Multifuel engines, which are characterized by compatibility with regard to the fuel used (alternatively operation e.g. with diesel fuels or kerosene, gasolines or diesel fuels and vegetable oils), no longer play a role today since the emission requirements for such engine concepts are no longer feasible.

References
[1] H. Oertel; M. Böhle; U. Dohrmann: Strömungsmechanik. 5th Ed., Vieweg+Teubner, 2008.
[2] R. van Basshuysen; F. Schäfer (Editors): Handbuch Verbrennungsmotor, 4th Ed., Vieweg+Teubner, 2007.
[3] K. Boulouchos: Strategies for Future Combustion Systems – Homogeneous or Stratified Charge? SAE 2000-01-0650.

Charge cycle and supercharging

Charge cycle

In combustion engines employing open processes and internal combustion, the charge cycle (exhaust and refill) system must serve two decisive functions:
- Replacement is employed to return the gas medium to its initial (start of cycle) condition,
- The oxygen required to burn the fuel is provided in the form of fresh air.

The charge cycle – also called gas exchange – can be assessed with the aid of the characteristic quantities defined in DIN 1940 [1]. For overall air flow (air expenditure $\lambda_a = m_g/m_t$) the entire charge transferred during the working cycle m_g is defined with reference to the theoretical maximum m_t for specific displacement. In contrast, the volumetric efficiency $\lambda_{a1} = m_z/m_t$ is based exclusively on the fresh charge m_z actually present or remaining in the cylinder. The difference between this and the total charge transfer m_g consists in the proportion of gas that flows directly into the exhaust tract in the overlap phase, making it unavailable for subsequent combustion.

The retention rate $\lambda_a = m_z/m_g$ is an index of the residual charge in the cylinder.

The scavenge efficiency $\lambda_S = m_z/(m_z+m_r)$ indicates the volume of the fresh charge m_z relative to the existing total charge, consisting of the fresh charge and the residual gas m_r. Here, the parameter m_r indicates the amount of residual gas from earlier working cycles remaining in the cylinder after the exhaust process.

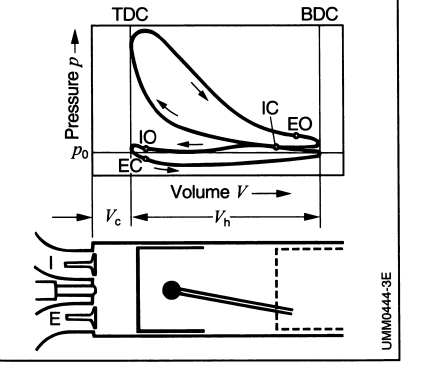
In a 2-stroke cycle, the gas is exchanged with every rotation of the crankshaft at the end of the expansion in the area around bottom dead center. In a 4-stroke cycle, separate intake and exhaust strokes provide a supplementary gas-exchange cycle.

4-stroke cycle

Valve timing – and thus gas exchange – are regulated by a control shaft (camshaft) rotating at half the frequency of the crankshaft by which it is driven. The camshaft opens the gas-exchange valves by depressing them against the valve springs to discharge the exhaust gas and to draw in the fresh gas (exhaust and intake valves respectively) (Figure 1). Just before piston bottom dead center (BDC), the exhaust valve opens and approx. 50 % of the combustion gases leave the combustion chamber under a supercritical pressure ratio during this predischarge phase. As it moves upward during the exhaust stroke, the piston sweeps nearly all of the combustion gases from the combustion chamber.

Shortly ahead of piston top dead center (TDC) and before the exhaust valve has closed, the intake valve opens. This crankshaft top dead center position is called the gas-exchange TDC (GTDC) or overlap TDC (OTDC) because the intake and exhaust processes overlap at this point) in order to distinguish it from the ignition TDC (ITDC). Shortly after gas-

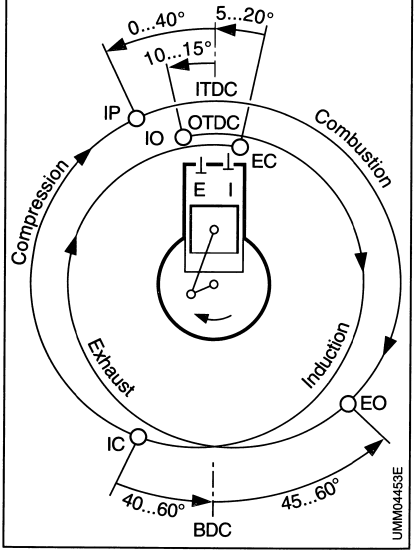
Figure 1: Representation of the 4-stroke gas-exchange process in the pV diagram
I Intake, E Exhaust,
BDC Bottom dead center, TDC Top dead center,
IC Intake closes, IO Intake opens,
EC Exhaust closes, EO Exhaust opens,
V_c Compression volume, V_h Displacement.
The arrows indicate the direction of the curves.



exchange TDC, the exhaust valve closes and, with the intake valve still open, the piston draws in fresh air on its downward stroke. This stroke of the gas-exchange process, the induction cycle, is completed shortly after bottom dead center (BDC) is reached. The subsequent two strokes in the 4-stroke process (Figure 2) are compression and combustion (expansion).

On throttle-controlled gasoline engines during the valve overlap period, exhaust gases flow directly from the combustion chamber into the intake passage, or from the exhaust passage back into the combustion chamber and from there into the intake passage. This internal exhaust-gas recirculation increases the combustion-chamber-charge temperature and the proportion of inert gas in the cylinder. A consequence of this is non-optimum power utilization in the upper load range. With fixed valve timing it is important to maintain a compromise, which is dependent above all in gasoline-engine applications on matching the engine operating strategy with the turbocharging unit.

Figure 2: 4-stroke gas-exchange process
E Exhaust, I Intake,
EO Exhaust opens, EC Exhaust closes,
IO Intake opens, IC Intake closes,
TDC Top dead center,
OTDC Overlap TDC, ITDC Ignition TDC,
BDC Bottom dead center,
IP Ignition point.



Early exhaust-valve timing allows sufficient time for the cylinder charge to escape, and thus guarantees low residual-gas compression as the piston sweeps through its upward stroke, although at the price of a reduction in the work index for the combustion gases.

The "intake valve closes" position (IC timing) exercises a decisive effect on the relationship between volumetric efficiency and engine speed. The maximum volumetric efficiency is achieved at low engine speeds when the intake valve closes early, and at high engine speeds when it closes late.

Obviously, fixed valve timing will always represent a compromise between two different design objectives: maximum brake mean effective pressure – and thus torque – at the most desirable points on the curve, and the highest possible peak output. The higher the engine speed at which maximum power occurs, and the wider the range of engine operating speeds, the less satisfactory will be the ultimate compromise. This tendency cannot be neutralized either by multi-valve cylinder heads with a larger intake-flow cross-sectional area.

At the same time, the demands for minimum exhaust emissions and maximum fuel economy mean that low idle speeds and high low-end torque (despite and along with high specific outputs for reasons of power-unit weight) are becoming increasingly important. These imperatives lead to the application of variable valve timing.

The advantages of the 4-stroke process are very good volumetric efficiency over the entire engine-speed range, low sensitivity to pressure losses in the exhaust system, and relatively good control of the speed-sensitive air expenditure through selection of appropriate valve timing and intake-system designs.

The disadvantages of the 4-stroke process are the high complexity of valve timing and the reduced power density of non-turbocharged 4-stroke engines because only every second crankshaft revolution is used for power output by combustion.

2-stroke cycle

To maintain gas exchange without an additional crankshaft rotation, the gases are exchanged in the two-stroke process at the end of expansion and at the beginning of the compression stroke. The intake and exhaust timing are usually controlled by the piston as it sweeps past the intake and exhaust ports in the cylinder housing near BDC (Figure 4). This configuration, however, requires symmetrical control times and involves the problem of directly flushing fresh mixture into the exhaust area (short-circuit scavenging). In addition, 15 to 25 % of the piston stroke cannot produce work because only charge volume V_f and not displacement volume V_h can be exploited for power generation (Figure 3).

As the two-stroke process lacks separate intake and exhaust strokes, the cylinder must be filled and scavenged using positive pressure, necessitating the use of scavenging pumps. In an especially simple and very frequently-used design, the bottom surface of the piston works in conjunction with a crankcase featuring a minimal dead volume to form a scavenging pump. Figure 4 shows a 2-stroke engine with crankcase scavenging and precompression along with the associated control processes. The processes which take place on the scavenging pump side are shown in the inner circle, while those occurring on the cylinder side are shown in the outer circle. The location of the intake and exhaust ports with the aim of maximum volumetric efficiency is also crucially dependent on the piston position. A hump-shaped increase in piston height can reduce direct throughflow from the intake area to the exhaust area.

The advantages of the 2-stroke process are high power output referred to engine weight and engine volume and a more uniform torque generation (one power stroke per revolution).

The disadvantages of the 2-stroke process are higher fuel consumption, lower mean pressures (due to poorer cylinder charge), higher thermal loads (due to the lack of a gas-exchange stroke), demanding mixture control, and higher HC emissions due to problematic cylinder scavenging. Two-stroke processes are no longer important due to the introduction of strict emission limits for most mobile applications.

Figure 3: Representation of the 2-stroke gas-exchange process in the pV diagram

p_0 Charge pressure, V_S Scavenging volume, V_c Compression volume, V_h Displacement, V_f Charge volume.

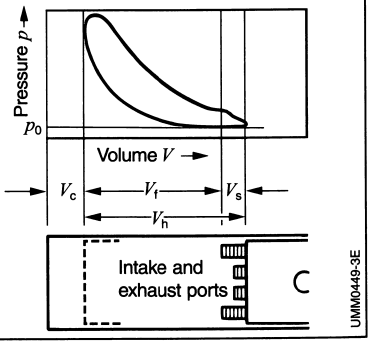
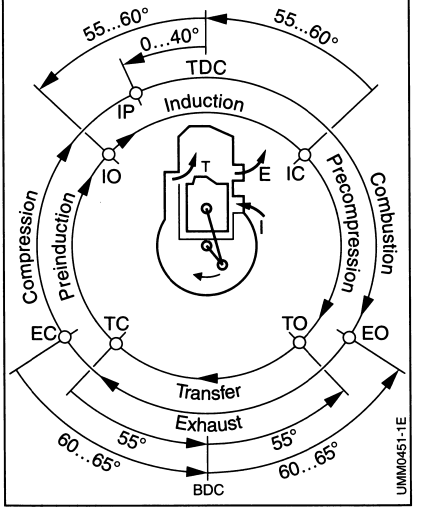


Figure 4: 2-stroke gas-exchange process with crankcase precompression

E Exhaust, I Intake, EO Exhaust opens, EC Exhaust closes, IO Intake opens, IC Intake closes, T Transfer passage, TO Transfer passage opens, TC Transfer passage closes, TDC Top dead center, BDC Bottom dead center, IP Ignition point.



Variable valve timing

Variable valve timing in a number of design variations is gaining acceptance with different objectives. The motivations behind the use of variable valve timing are as follows: increased power and torque, dethrottling, control of the residual-exhaust-gas content, cylinder shutoff, charge movement, cold-starting and cold-running performance, control-engineering optimization (e.g. charge mass, residual-exhaust-gas content), influencing the rotational speed of the exhaust-gas turbocharger, and exhaust-gas temperature-management measures.

Today, different systems are used to provide variable valve timing. There are different motivations behind variable valve timing, depending on the combustion process and engine concept (homogeneous- or stratified-charge operation, diesel).

Camshaft adjustment

The concept of adjusting the intake and exhaust camshafts is increasingly becoming an accepted measure, offering a valuable degree of variability in the entire program map. This concept provides the option of adjusting the phase of the camshafts with respect to the crankshaft and thereby shifting the valve timing while maintaining the same form of opening. Electric or electrohydraulic activation of the camshaft adjuster is the state-of-the-art technology used.

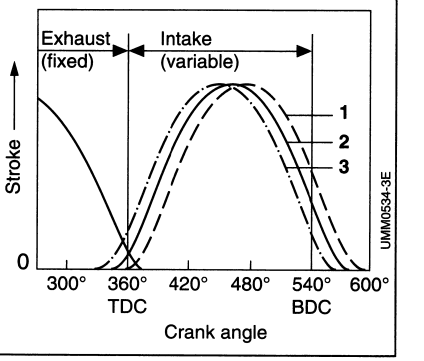
In homogeneously operated gasoline engines the intake camshaft is subject to retarded adjustment for example in the low speed and load range while at the same time the exhaust valve is subject to very advanced closing (Figure 5). This minimal valve overlap delivers a minimal throughflow of fresh-air mixture into the exhaust area. At the same time the highly retarded opening of the intake valve also causes dethrottling by the associated retarded closing of the intake valve, because the cylinder charge is again discharged by the intake valve after bottom dead center. A rapid torque increase can be achieved for a desired load point by advancing the intake camshaft.

Exhaust-camshaft adjustment also provides for further degrees of freedom. Optimal adjustment of the intake and exhaust camshafts is dependent on numerous factors. The operating mode is crucial as well as the position in the engine map. For example, a significant factor is whether exhaust-gas turbocharging is used or whether lean-burn operation (diesel engine, gasoline-engine stratified-charge process, etc.) is applied. One example of the beneficial combination of camshaft adjustment with an exhaust-gas turbocharger is the map-dependent maximum valve overlap (very retarded closing of the exhaust valve and very advanced closing of the intake valve), at which a large proportion of mixture flows directly from the intake to the exhaust area. This can give rise to the beneficial effect of an increase in air-mass flow by the exhaust-gas turbocharger together with an advantageous increase in turbocharger speed. Rotating the intake and exhaust camshafts offers optimization opportunities for a large number of questions and is therefore being increasingly used at least in modern gasoline engines.

Systems with cam-shape control mechanism

As well as continuous phase adjustment, simpler systems are also available which alternate digitally between individual cam shapes. Typically, such systems adjust a very large cam and valve lift for ranges

Figure 5: Intake-camshaft rotation
1 Retard, 2 Standard, 3 Advance.



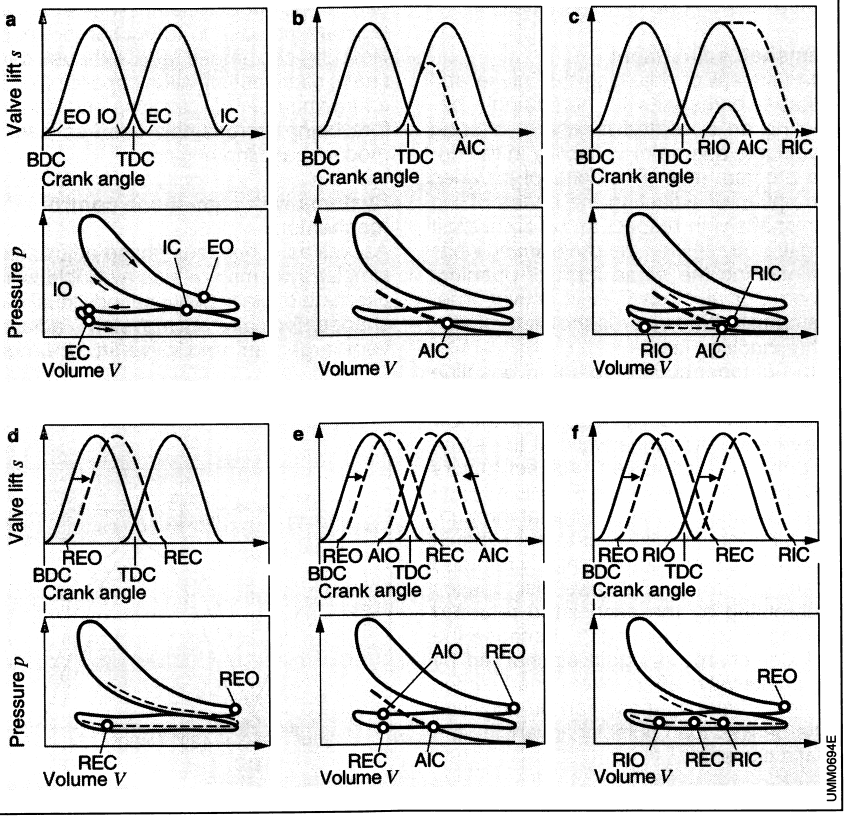
close to full load and a small cam and valve lift for dethrottling for part-load ranges. This is achieved for the most part by integrating a control mechanism on a rocker arm or bucket tappet, which thereby activates one or more (typically two) available cams on the camshaft.

Systems that can alternate between two camshaft rotation positions are also used. In contrast to fully variable systems, these systems omit an exact positioning command to the camshaft phase adjuster.

Other variable systems

Increasingly, fully variable systems for use in gasoline engines are being developed and are even already appearing in production vehicles. These systems focus on optimizing the charge-cycle phase. There are different successful approaches to suit the engine configuration and operating point, each approach optimizing charge-cycle losses, the residual-exhaust-gas content or engine power (Figure 6).

Figure 6: Influence of valve timing on the charge cycle
a) Conventional valve timing, b) Advanced "Intake closes",
c) Valve-lift reduction and retarded "Intake closes", d) Retarded "Exhaust closes",
e) Retarded "Exhaust closes" and advanced "Intake closes",
f) Retarded "Exhaust closes" and retarded "Intake closes".
BDC Bottom dead center, TDC Top dead center, IO Intake opens,
EO Exhaust opens, IC Intake closes, EC Exhaust closes,
A Advance, R Retard.



Mechanical, electromechanical, electrohydraulic or electropneumatic adjustment concepts can be used to provide a fully variable valve gear.

Mechanical systems
Mechanical, fully variable valve-gear systems typically consist of a combination of an adjustment mechanism, which permits variation of the valve lift, and a camshaft phase adjuster. The key function is the provision of variable cam lift.

A suitable design solution is provided in this case by a conventional camshaft, which does not however act directly on the valve via a bucket tappet or a rocker arm. The cam lift is adjusted by an intermediate lever whose fulcrum can be varied by means of an eccentric shaft (Figure 7 and Figure 8). Electric DC motors are typically used to drive such arrangements. BMW Valvetronic is an example of a fully variable mechanical valve gear [2]. Further advantages of a fully variable valve gear are the optimization of uniform charge distribution and the potential to increase charge movement by having the two intake valves opening at different times in the lower load range thanks to slightly differently ground eccentricities.

Figure 7: Valve-lift curves of a mechanical fully variable valve gear

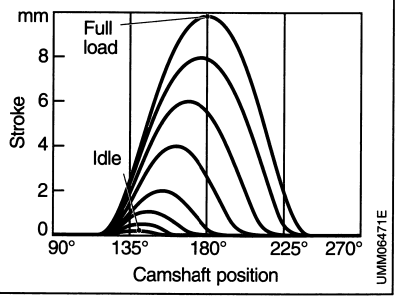
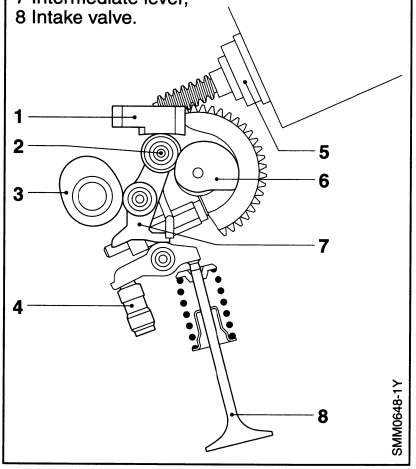
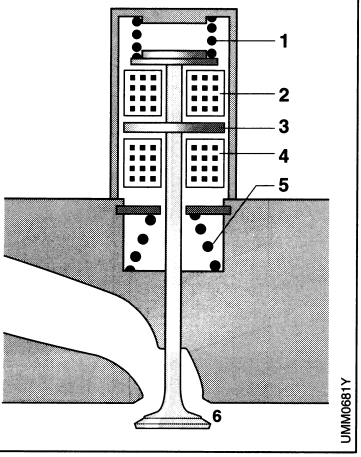


Figure 8: Basic principle of a mechanical fully variable valve gear (BMW Valvetronic)
1 Guide block,
2 Center of rotation,
3 Camshaft,
4 Hydraulic valve-clearance compensation,
5 Servo unit with spiral-toothed gear,
6 Eccentric shaft,
7 Intermediate lever,
8 Intake valve.



Electromechanical systems
Electromechanical systems (electromechanical valve gear) are still in the development stage. In an electromechanical valve gear, electrically actuated magnets are used as actuators for valve timing (Figure 9, see also [3]). The high power

Figure 9: Electromechanical systems
1 Compression spring,
2 Closing magnet,
3 Armature,
4 Opening magnet,
5 Compression spring,
6 Intake valve.



demand of the actuators must be noted in particular. The system, comprising valve, magnet and coil, is made to resonate in order to minimize this demand. Because of the high power demand and the complexity involved, electromagnetic valve gears have not yet made it to series production.

Electrohydraulic systems

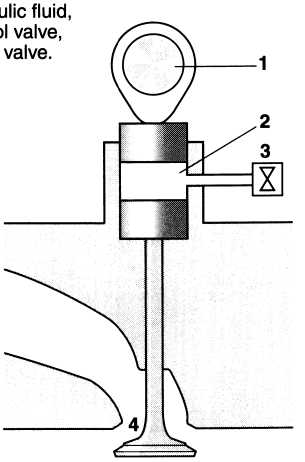
Electrohydraulic systems (electrohydraulic valve timing) are an alternative to the fully variable mechanical valve gear. Different principles are applicable here.

The principle of "lost motion" is an effective approach. A motion is stipulated by a hydraulic intermediate element via a camshaft (Figure 10). An electrically actuated hydraulic valve provides the opportunity of not transferring completely the motion stipulated by the cam. The cam shape therefore stipulates an envelope (Figure 11).

An alternative is presented by a system which directly actuates the gas-exchange valves via a hydraulic pressure accumulator and electronically controlled hydraulic valves. Systems of this type are still in the development stage and are not yet in series production.

Figure 10: Working principle of electrohydraulic systems

1 Camshaft,
2 Hydraulic fluid,
3 Control valve,
4 Intake valve.



Depending on the point and duration of activation, a proportion of the hydraulic fluid escapes through the hydraulic valve. Hydraulic power loss is therefore a drawback of this system.

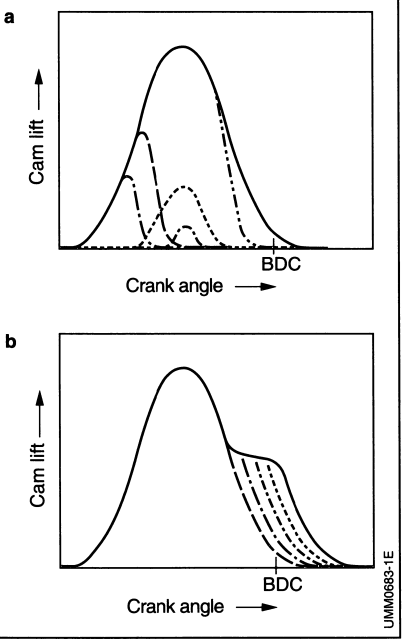
Electrohydraulic systems have been deployed since 2004 in Caterpillar commercial-vehicle systems and since 2010 in Fiat's MultiAir system.

Electropneumatic systems

Electropneumatic systems are still in the development stage. The use of such systems in large-scale production is not anticipated at present. As well as exhibiting complex control characteristics, these systems above all demonstrate unfavorable pneumatic drive power. Current systems are used for research purposes. The power input of the compressed-air supply must be noted and necessarily included in an efficiency balance.

Figure 11: Cam lifts of electrohydraulic systems for different applications

a) Advanced closing of intake valve,
b) Variable retarded closing of intake valve.



Supercharging processes

The power of an engine is proportional to its air-mass flow. As this air throughput, in turn, is linearly dependent on air density, the power of an engine, given a specific displacement and engine speed, can be increased by compressing the air before it enters the cylinders, i.e. by supercharging. The supercharging ratio indicates the density rise as compared to a naturally aspirated engine.

Thermodynamically speaking, the best results would be delivered by isothermal compression, but this cannot be technically achieved. Reversibly adiabatic compression in the supercharger unit serves as the ideal comparison process for the actual sequences; in practical terms, increased density is accompanied by losses.

The extent of the supercharging ratio is limited in gasoline engines by the occurrence of knocking combustion and in diesel engines by the maximum permissible peak pressure in the cylinder. To offset this limitation, supercharged engines usually have lower compression ratios than their naturally aspirated counterparts.

Dynamic supercharging

The gas-exchange processes are not only influenced by the valve timing, but also by the geometry of the intake and exhaust lines. Induced by the induction work of the piston, as the intake valve opens it initiates a suction wave in the intake manifold that reflects off the open end of the intake manifold and returns as a pressure wave to the intake valve. These pressure waves can be used to increase the air-mass intake (Figure 12). In addition to the geometry in the intake manifold, the impact of this supercharging effect based on gas dynamics also depends on the engine speed (see Ram-tube supercharging, Tuned-intake-tube charging, Variable-geometry intake-manifold systems).

Mechanical supercharging

In mechanical supercharging, the supercharger is driven directly by the internal-combustion engine (see Turbochargers and superchargers). The charger and the internal-combustion engine are mechanically coupled to each other. Types of mechanical supercharger unit are positive-displacement superchargers (compressors) of various designs (e.g. Roots superchargers, spiral-type superchargers) and hydrokinetic flow compressors (e.g. radial compressors).

The crankshaft and charger shaft feature a fixed transmission ratio in system designs used up to now. Mechanical or electromagnetic clutches can be used to control supercharger activation. The boost pressure is generally adjusted by a bypass device with control flap (wastegate).

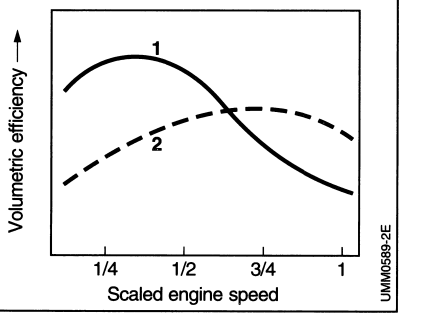
Advantages of mechanical supercharging

- The supercharger is installed on the cold side of the engine.
- The engine's exhaust system is not influenced by supercharger components.
- The supercharger responds almost immediately to load changes.

Disadvantages of mechanical supercharging

- The drive power must be branched from the effective power of the engine, which results in increased fuel consumption.

Figure 12: Increase of volumetric efficiency through dynamic supercharging
The engine speed is scaled to the nominal speed.
1 System with dynamic supercharging,
2 System with standard intake manifold.



- Low-noise performance can only be achieved by additional measures.
- Comparatively high structural volume and weight.
- The supercharger unit has to be positioned at the engine's belt level.

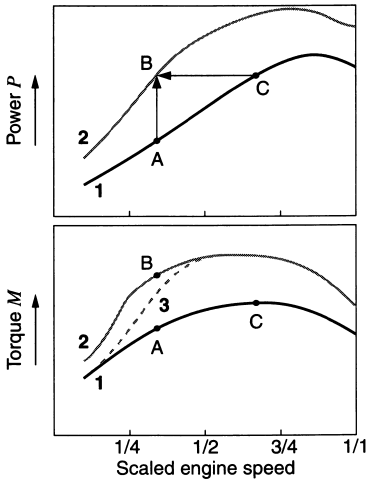
Exhaust-gas turbocharging

In exhaust-gas turbocharging, the power for driving the charger is extracted from the engine's exhaust gas, i.e. some of the energy contained in the exhaust gas is converted into mechanical energy by an exhaust-gas turbine. In this way, the process exploits some of the enthalpy that remains unused (owing to crankshaft-assembly expansion limits) by naturally aspirated engines. Exhaust backpressure however increases in the process. Hydrokinetic flow compressors are exclusively used to compress the induction air (see Turbochargers and superchargers).

Exhaust-gas turbochargers are usually designed to generate a high boost pressure even at low engine speeds. In other words, the turbocharger turbine is generally designed for an engine mid-speed range. However, without further measures, boost pressure at the high end of the engine speed range could increase to levels that would place excessive load on the engine. The turbine is therefore equipped with a bypass valve which, from a specific operating point, diverts some of the exhaust-gas mass flow past the turbine. However, the energy of this exhaust gas remains unutilized. Much more satisfactory results can be achieved with a compromise between high boost pressure in the low engine-speed range and avoidance of engine overload at the high end of the engine-speed range by employing a turbocharger with variable turbine geometry (VTG). The adjusting blades used for this purpose adapt the flow cross-section and the angle of impact to the turbine (and thus the exhaust-gas pressure applied to the turbine) by varying the position of the guide blades (see Turbochargers and superchargers).

Figure 13: Comparison of power and torque curves between naturally aspirated engine and turbocharged engine

The engine speed is scaled to the nominal speed.
1 Naturally aspirated engine in stationary operation,
2 Turbocharged engine in stationary operation,
3 Torque build-up of a turbocharged engine in transient operation.
A→B Higher torque and higher power at identical speed for turbocharged engine.
C→B Identical power for turbocharged engine at lower speed.



Advantages of exhaust-gas turbocharging

- Considerable increase in power output per liter.
- Improvement in fuel-consumption figures relative to naturally aspirated engines with the same output power.
- Improvement in exhaust-gas emission rates.
- Comparatively small structural volume.
- Can be used with low-pressure-side exhaust-gas recirculation.

Disadvantages of exhaust-gas turbocharging

- Installation of the turbine side of the turbocharger in the hot exhaust-gas system requiring materials resistant to high temperatures.
- Increased thermal inertia in the exhaust-gas system.
- Without further measures, comparatively low starting torque from low-displacement engines.

Special forms of supercharging

In electrically assisted exhaust-gas turbocharging systems, an additional electric motor serves to maintain the turbocharger at a high speed in the absence of exhaust-gas mass flow. The produces benefits above all in transient vehicle operation and at low engine speeds, but the complexity and the electric power consumption are very high. The drawbacks of the exhaust-gas turbocharger due to inertia can be reduced with electrically assisted exhaust-gas turbocharging.

A special but only theoretical form of supercharging is pressure-wave supercharging (Comprex); this however failed with regard to production maturity. The working principle is based on the reflection properties of pressure waves which propagate in a rotating cell rotor (see Turbochargers and superchargers). The chief advantage is the transient response, which facilitates a very rapid torque build-up. However, the costs involved with pressure-wave superchargers are high and the required drive necessitates additional packaging measures.

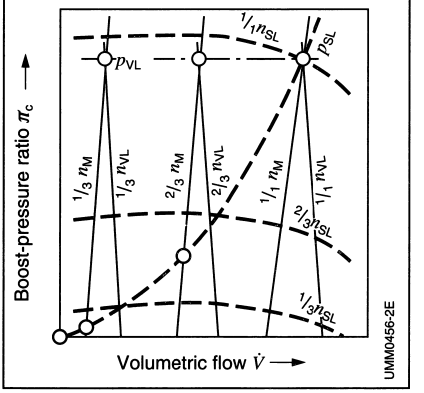
Volumetric-flow map

A clear illustration of the relationship between the engine and the supercharger is provided by the diagram for pressure vs. volumetric flow (Figure 14), in which the pressure ratio π_c of the supercharger is plotted against volumetric flow $\dot{V}$.

The curves for unthrottled 4-stroke engines (diesel) are particularly descriptive because they contain sloped straight lines (engine mass-flow characteristics) which represent increasing air-throughput values as the pressure ratio $\pi_c = p_2/p_1$ (where p_1 denotes the ambient pressure and p_2 the charge-air pressure) increases at constant engine speed.

Figure 14: Pressure vs. volumetric-flow map of mechanically driven positive-displacement supercharger and hydrokinetic flow compressor

n Speed, p_L Boost pressure.
Subscripts:
VL Positive-displacement supercharger,
SL Hydrokinetic flow compressor,
M Engine.



The diagram shows pressure ratios which result at corresponding constant supercharger speeds for a positive-displacement supercharger and a hydrokinetic compressor.

Only superchargers whose delivery rates vary linearly with their rotational speeds are suitable for vehicle engines. These are positive-displacement superchargers of piston or rotating-vane design or Roots blowers. Mechanically driven hydrokinetic flow compressors are not suitable.

Exhaust-gas recirculation

Operating principle

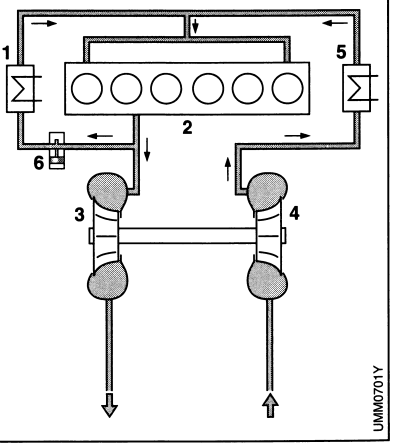
With exhaust-gas recirculation (EGR) hot exhaust gas is removed, cooled in an exhaust-gas cooler, mixed with fresh air, and returned to the combustion chamber. The exhaust-gas recirculation rate is adjusted by a valve. It is used in diesel engines and also increasingly in gasoline engines.

Exhaust-gas recirculation in diesel engines

In diesel engines the untreated emissions of NO_x are reduced significantly by exhaust-gas recirculation. Essentially this is done by lowering the peak temperature in the cylinder, which is obtained by the fact that the on the one hand the recirculated exhaust gas reduces the combustion velocity and on the other hand the exhaust gas on account of its now higher heat capacity – above all of the CO2 – for the same supplied energy does not become as hot as exhaust gas without exhaust-gas recirculation.

Figure 15: High-pressure exhaust-gas recirculation

The arrows denote the flow direction.
1 EGR cooler for high-pressure EGR,
2 Engine,
3 Exhaust-gas-turbocharger turbine,
4 Exhaust-gas turbocharger compressor,
5 Intercooler,
6 EGR valve.



Exhaust-gas recirculation in gasoline engines

In gasoline engines the idea is not to reduce the nitrogen oxides. The aim is primarily to reduce the exhaust-gas temperature to protect components (catalytic converter and turbocharger). This is very effective particularly in engine full-load operation because the otherwise usual process of enriching the air/fuel mixture to reduce the exhaust-gas temperature can be partially or completely dispensed with. This results in a dramatic improvement in fuel consumption.

An additional consideration in gasoline engines the possibility of diluting the charge by means of exhaust-gas recirculation without compromising the function of the three-way catalytic converter. Engine dethrottling can be achieved by increasing the EGR rate particularly in engine part-load operation, which also results in improved fuel efficiency.

Exhaust-gas recirculation variants

Essentially, one of two variants – high-pressure and low-pressure recirculation – is used, depending on whether the exhaust-gas cooler is located upstream or downstream of the turbocharger. Whereas high-pressure recirculation is found primarily in commercial vehicles, both variants are used in passenger cars.

High-pressure EGR

With high-pressure recirculation (Figure 15), exhaust gas is recirculated on the high-pressure side, i.e. upstream of the turbocharger turbine. It has the advantage that the pressure difference required to transport the exhaust gas from the exhaust side to the intake side is higher and can even be altered by means of variabilities at the turbocharger (wastegate or variable turbine geometry). High EGR rates can basically be achieved because the EGR section is short and can therefore be filled and emptied quickly. Thanks to this dynamic advantage, the EGR system can respond relatively quickly to changing operating conditions.

Because the recirculated exhaust gas is not mixed with fresh air until downstream of the intercooler, the intercooler and the turbocharger compressor are not

permeated and thus not fouled by the exhaust gas. The performance of these components therefore remains constant over their full service lives. By contrast, significantly greater demands are placed on the EGR cooler on the high-pressure side. This can be put down to the higher pressures and temperatures during operation and to the already mentioned component sooting.

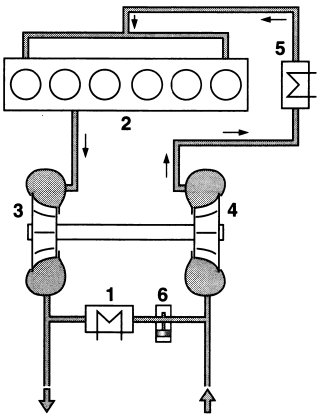
Low-pressure EGR

This form of exhaust-gas recirculation routes the exhaust gas only downstream of the turbine through the EGR cooler and mixes it with the inducted fresh air upstream of the turbocharger compressor (Figure 16). The exhaust gas arrives at the EGR cooler only after it has passed through the diesel particulate filter (DPF, not shown in the illustration) and the soot has been removed. Thus the EGR cooler is not fouled and performs consistently over its full service life. Because of the lower temperature in this area of the exhaust system the temperature stresses in the EGR cooler are also much lower than is the case for high-pressure recirculation. The lack of fouling together with a lower inlet temperature provide for a lower cooler outlet temperature than with high-pressure recirculation. The reduction in NO_x emissions that can thereby be achieved can be sufficient in many cases to eliminate the need for further exhaust-gas treatment, such as the SCR system for example.

One drawback of low-pressure recirculation is the reduced dynamic response. The system is relatively far removed from the engine and, due to the long distances involved, cannot adapt the exhaust-gas recirculation rate quickly. It is necessary to set the EGR rates slightly lower than for high-pressure recirculation so as to prevent the engine from being "flooded" with exhaust gas.

Figure 16: Low-pressure exhaust-gas recirculation

The arrows denote the flow direction.
1 EGR cooler for low-pressure EGR,
2 Engine,
3 Exhaust-gas-turbocharger turbine,
4 Exhaust-gas turbocharger compressor,
5 Intercooler,
6 EGR valve.



References

[1] DIN 1940: Reciprocating internal combustion engines; terms, formulae, units.

[2] R. Flierl; R. Hofmann; C. Landerl; T. Melcher; H. Steyer: Der neue BMW-Vierzylindermotor mit Valvetronic. MTZ 62 (2001), Volume 6.

[3] P. Langen; R. Cosfeld; A. Grudno; K. Reif: Der Elektromechanische Ventiltrieb als Basis zukünftiger Ottomotorkonzepte. 21st International Vienna Motor Symposium, Progress Reports VDI, Series 12, No. 420, Volume 2, 2000.

[4] R. van Basshuysen; F. Schäfer (Editors): Handbuch Verbrennungsmotor, 5th Edition, Vieweg-Verlag, 2009.

Reciprocating-piston engine

Components

In addition to the crankshaft drive, which makes a vital contribution by generating torque, the significance of the cylinder head is crucial to the efficiency of an internal-combustion engine (Table 1).

Crankshaft drive and crankcase
The crankshaft drive comprises the following components: piston, piston rings, connecting rod, and crankshaft. What all the crankshaft-drive components have in common is a translational and rotational motion. The tribological design of the crankshaft drive is already hugely significant in view of its robustness and reduced engine friction.

Piston
Because of the mechanical and thermal loads to which it is subjected, the piston is a highly complex component in an internal-combustion engine. The recess and piston-surface geometries have a crucial impact on mixture formation and combustion. In addition, the primary function of the piston is to transmit force to the connecting rod. Complex mechanical stresses occur in the process with simultaneously very high local temperatures up to and in excess of 300 °C.

The role of the piston recess is even more important in diesel engines. With the high compression ratio and a typically flat cylinder head, the entire combustion chamber must be accommodated in the piston recess. Various piston-recess designs are successfully used. Figure 1 shows examples of the design variations used in diesel engines.

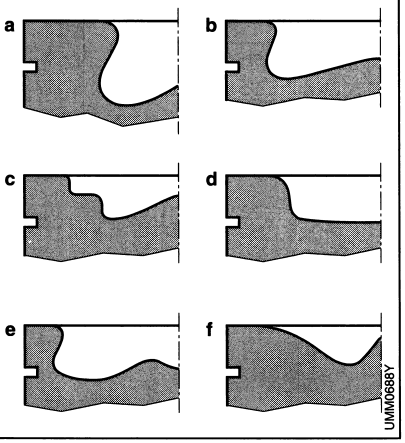
The piston recess decisively defines the combustion-chamber shape and at the same time influences the strength of the piston crown, a factor which is vital to the function of force input into the piston pin. Especially when it comes to aluminum pistons, which are currently the preferred choice mainly for weight reasons, the requirements of high heat resistance, low weight and cylinder peak pressure call for technically sophisticated solutions. Brass bushings are used to accommodate the piston pin in diesel engines subject to high loads.

Diesel engines can also utilize steel and even gray-cast-iron pistons as well as aluminum designs. Steel pistons are used in applications with peak pressures well in excess of 200 bar. An important distinguishing feature of steel is its lower heat conduction.

Modern turbocharged engines with high cylinder peak pressures and high power outputs also require a piston spray injector with an additional piston cooling duct for cooling purposes (Figure 2). Here engine oil is sprayed through a vertically arranged piston spray injector into the piston cooling duct. The inlet and outlets for the piston cooling duct are located on the underside of the piston.

Piston rings form the sealing element between the combustion chamber and the crankcase (Figure 3). As well as providing the seal, the piston rings also perform the important tasks of conducting heat to the

Figure 1: Piston shapes for diesel engines
a) Deep omega recess,
b) Flat omega recess,
c) Stepped recess,
d) Dish recess,
e) Eccentric recess design,
f) W recess.



cylinder wall and controlling the entry of oil into the combustion chamber.

Three rings are typically used, the top ring acting as a compression ring and the bottom ring as an oil control ring. The middle ring frequently performs both functions. Because particulate filters become clogged with ash, which is also a product of burnt oil in the combustion chamber, oil consumption and therefore the coordination of the piston-ring set and honing are vitally important.

Table 1: Components of a 4-stroke reciprocating-piston engine

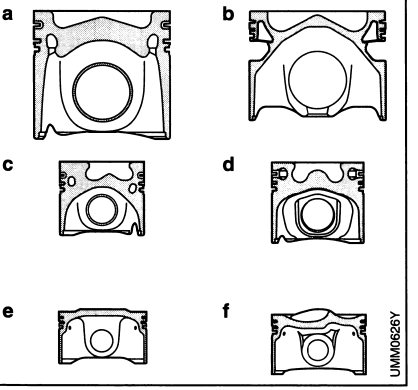
Assembly	Component	Purpose and function	Stress
Crankcase	Crankcase	– Housing for crankshaft drive – Accommodation of cylinder head and counterpart for accommodation of cylinder-head bolts – Accommodation of crankshaft bearings – Accommodation of accessories	Deformation – flow of force from cylinder-head bolts to crankshaft, bending moments, vibrations, oscillations and natural frequencies
	Bearing surface	Different approaches to structural solution of bearing surface: – Accommodation of liners with seal (wet liners) – Accommodation of cast liners (dry liners) – Piston bearing surface directly integrated in crankcase	Deformation (out-of-round), tribological counterpart to piston
Crankshaft drive	Piston	– Conversion of energy in combustion chamber (gas pressure) into a translational motion – Dissipation of heat from combustion chamber to cooling medium – Support of the normal forces occurring as a result of the connecting-rod inclination	Complex mechanical and local thermal stress above 350 °C in the recess area
	Piston rings	– Sealing of combustion chamber in relation to crankcase – Avoidance of oil entry into cylinder – Dissipation of heat to cylinder wall	High mechanical bending stress and high thermal stress, complex tribology
	Connecting rod and pin	– Absorption of the forces acting on the piston and transmission to the crankshaft for torque build-up	High mechanical compressive, tensile and bending stresses, complex tribological processes
	Crankshaft	– Absorption of the forces acting on the connecting rod and conversion of oscillatory to rotational motion. Thus build-up of a torque for transmission to the drivetrain	High mechanical compressive, tensile and bending stresses (buckling), complex tribological processes
Cylinder head	Cylinder head	– Control of charge cycle – Cooling of areas close to the combustion chamber – Accommodation of injectors (in diesel engine and gasoline engine with direct injection) and of spark plug (in gasoline engine)	High stress due to thermomechanical fatigue behavior and cold-hot expansion

Table 1: Components of a 4-stroke reciprocating-piston engine (continued)

Assembly	Component	Purpose and function	Stress
Cylinder head	Camshaft (typically in cylinder head)	– Conversion of a rotational motion into a translational motion to open and close the valves	Hertzian stress
	Valve control	– Conversion of a rotational motion into a translational motion to open and close the valves	High valve-closing speeds require large valve spring forces
	Valves	– Opening of passages to intake and exhaust ports – Sealing of combustion chamber during compression phase	Valve-seat wear due to high valve-closing speeds, thermal load at exhaust valve
	Ports/ passages	– Supply and discharge of cylinder charge – Provision of the desired cylinder-charge movement	Local thermal load
	Base plate	– Accommodation of combustion-chamber pressure and sealing	Local thermal load and bending
	Geartrain	– Driving of camshaft by belt, chain or gear wheels	Rotational irregularities with alternating contact flanks
Other components	Oil pump	– Provision of the required oil delivery rate and provisions of the required oil pressure	Torque particularly with viscous oil
	Oil filter	– Filtering of contaminants from engine oil	–
	Water pump	– Provision of the required cooling-medium delivery rate – Provision of the required cooling-medium mass flow	Cavitation

Figure 2: Piston shapes in various engine designs

a) Commercial-vehicle aluminum diesel-engine piston with ring carrier and cooling channel,
b) Commercial-vehicle forged-steel piston,
c) Passenger-car aluminum diesel-engine piston with ring carrier and cooling channel,
d) Passenger-car aluminum diesel-engine piston with cooled ring carrier,
e) Passenger-car aluminum piston for MPI gasoline engine (Multi-Point Injection),
f) Passenger-car aluminum piston for gasoline engine with GDI (gasoline direct injection).



Connecting rod

The connecting rod (Figure 4) links the piston with the crankshaft. It transmits the piston's gas and inertial forces introduced into the piston pin to the crank pin. It is therefore subjected to tensile, compressive and bending stresses. It also carries the bearings for the crankshaft and usually for the piston pin. It must therefore be particularly stiff in its design, especially in the connecting-rod eyes, which accommodate the bearings. For this reason, it is usually designed with a double-T-shaped shank profile and drop-forged with tempering steels. Sintered materials can also be successfully used. Annealed cast iron can also sometimes be used in low-load applications (gasoline engines). Rarely encountered is the combination of a standard connecting rod and a forked connecting rod in V-engines which then permits an exactly opposed cylinder arrangement.

The length of the connecting rod is determined by the piston stroke and the counterweight radius (the geometric path of the outer contour is called the "connecting-rod violin"). In commercial-vehicle engines with a relatively large stroke/bore ratio of more than 1.1 and large bearing diameters an angled split of the big connecting-rod eye is customary. This enables the piston to be removed without the crankshaft having to be removed also. However, this angled split generates considerable shear forces even when the connecting rod is positioned vertically at top dead center. These forces are usually absorbed by toothing, groove and spring or by cracked surfaces. Cracking of the big connecting-rod eye is achieved by a "fracture separation" using extremely high (hydraulic) expansion forces and a targeted weakening of the cross-section at the big connecting-rod eye. The big-end bearing cap and the connecting rod can only be installed as a single unit.

The bolts are designed as through-bolts with or without nuts. Adapter sleeves or bolt adapter collars ensure that the bearing shells are aligned when seated.

Lateral guidance of the connecting rod can be performed in the crankshaft at the big connecting-rod eye or in the piston at the small connecting-rod eye.

The small connecting-rod eye must absorb considerable forces and is therefore often form-bored in order to avoid misalignment. For diesel engines with high peak pressures in particular and in accordance with high gas forces, a special

Figure 3: Piston-ring shapes and configurations

Diesel engine:
1 Keystone ring, crowned,
2 Taper-face compression ring with inside bevel,
3 Stepped compression ring,
4 Double-beveled ventilated oil control ring with spiral-type expander.

Gasoline engine:
5 Plain compression ring, crowned,
6 Taper-face compression ring,
7 Stepped ring,
8 Double-beveled ring,
9 Multipart steel oil ring.

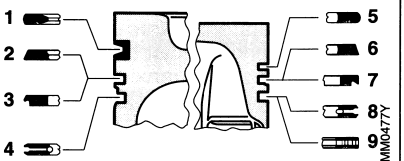
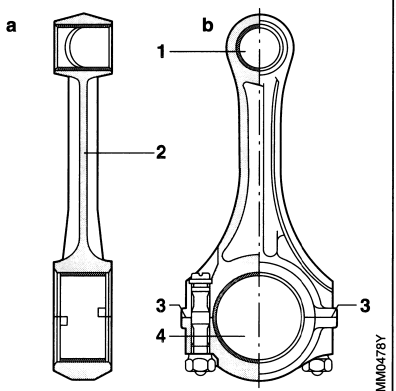


Figure 4: Connecting rod for a passenger-car engine

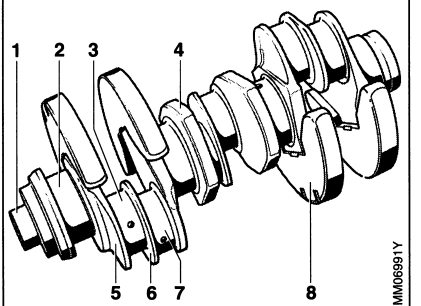
a) Sectional view, b) Side view.
1 Small connecting-rod eye,
2 T-shaped shank profile,
3 Straight split,
4 Big connecting-rod eye.



shape is chosen for the small connecting-rod eye which ensures a high contact surface on the pin without weakening the piston. This may be trapezoidal (narrow at the top, wide at the bottom) or stepped (narrow shoulder at the top, wide at the bottom).

The connecting rod is designed in line with the calculation of the gas and inertial forces (see Crankshaft-assembly design). The case of buckling is also calculated. Both buckling in the plane perpendicular to the crankshaft as a hinged beam and buckling in the crankshaft-axis plane (engine longitudinal plane) as a fixed beam must be checked. The critical cross-sections in the transition from the dimensionally stable connecting-rod eye to the double-T shank are usually recalculated with the aid of FEM simulations (finite-element method). The loads caused by the gripping of the bearing caps must not be underestimated and ignored.

Figure 5: Crankshaft for a commercial-vehicle engine
1 Front end of crankshaft,
2 Main bearing,
3 Connecting-rod bearing journal with oil bore,
4 Main web (without counterweight),
5 Main webs,
6 Intermediate web,
7 Offset connecting-rod bearing journal (split-pin),
8 Counterweight (bolted).



Crankshaft

The crankshaft with its throws (Figure 5) converts the oscillating motion of the piston by way of the connecting rod into a rotatory motion. The translational energy is transformed into a rotating torque, which can be picked off by a coupling.

The main bearings of the crankshaft are fixed in the crankcase, while the crank pins in the throws run in the bearings fixed to the connecting rods. The counterparts to the throws with crank pins are the counterweights, which are either cast, forged or bolted on. Depending on the engine type, the crankshaft main bearings are arranged after each cylinder – particularly in in-line engines and diesel and gasoline engines subject to high loads – or only after every second throw (gasoline engines subject to low loads). V-engines usually feature one crank pin for two adjacent connecting rods on one throw. This also applies to forked connecting rods, which permit two exactly opposing cylinder banks. In order to facilitate a uniform ignition interval in V6 engines with a 90° bank angle, split-pin designs which permit a 30° offset of two crank pins in a very tight space without a main bearing in between are increasingly being used. This form of design is used even in commercial-vehicle diesel engines subject to high loads.

Crankshafts are usually forged (drop forgings or open-die forgings, depending on the size) and, in lower-load applications, increasingly cast (nodular graphite iron). Built-up crankshafts are also used in large engines. Bores for supplying lube oil to the connecting-rod bearings are incorporated in such a way as not to generate further increases in stress.

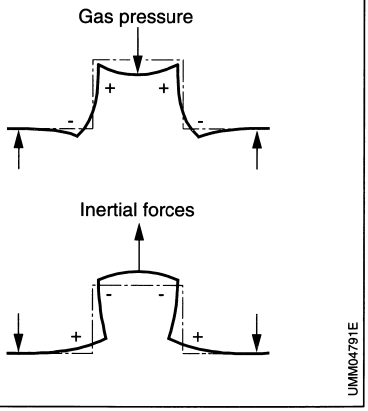
The gas and inertial forces (see Crankshaft-assembly design) subject the crankshaft to bending stress in particular (Figure 6). If a main bearing is incorporated after each throw, the bending loads on the crankshaft are low on account of tangential forces. However, particular attention must be paid to the transition between crank pin and web because this is where the greatest stresses occur. Because of the large number of main bearings, the clamping of the crankshaft represents theoreti-

cally a statically overdefined system – in the simulation, however, a statically defined state which correctly describes the critical state is assumed.

Another important factor is the torsional vibrations, which, on account of the alternating forces of each individual cylinder (torsional force, see Crankshaft-assembly design) excite the overall crankshaft-assembly system and, in view of the not inconsiderable length of the crankshaft, can approach the crankshaft's natural resonance.

The oscillatory system – consisting of piston, connecting rod and crankshaft – is a complex system of constantly changing torsional forces. The further the counterweights are arranged to the outside, the higher the excitation. For this reason, heavy-metal counterweights are also bolted on in some cases. A simple simulation can be performed by assuming a smooth, inertialess and elastic shaft with attached equivalent masses. The oscillation-equation system can then be set up with the aid of the torsion-spring rate of the torque shaft. Cast shafts are ascribed a certain level of intrinsic damping. The crankshaft is thus subjected to both a locally different and a time-varying torsional load.

Figure 6: Crankshaft throw
Primary stresses and deformations due to gas pressure and inertial forces.



The torsional vibrations of the crankshaft, which are also called 2nd-order torsional vibrations (see Crankshaft vibration damper) must be reduced by suitable vibration dampers. This is particularly necessary to protect the crankshaft against excessively high loads.

Crankcase

The crankcase supports the force-transfer mechanism between cylinder head and crankshaft assembly. It bears the crankshaft assembly's support bearings, and incorporates (or holds) the cylinder sleeves. Also included in the block are a separate water jacket and sealed oil chambers and galleries. Furthermore, it must absorb the torque reaction forces and provide locations for mounting many add-on components.

A rigid framework-like structure (support-beam concept) is normally used to absorb the forces; this structure is designed to ensure a direct, linear and bending-moment-free flow of force. This is usually achieved by walls with reinforcing ribs and in the case of the bolt-through design by through-bolts.

The cylinder head is usually bolted as a separate component – in larger engines also as an individual cylinder head – to the crankcase. Often four bolts per head are used, with adjacent cylinders usually sharing two bolts. Six or eight bolts may also be used in engines with high peak pressures, typically above 200 bar. The purpose of this in particular is to distribute as evenly as possible the pressure force of the cylinder-head gasket.

The crankshaft-bearing caps are usually bolted from underneath to the crankcase. In the bolt-through design a long bolt performs this function from the head to the bearing cap. The advantage of this design is that it permits the use of high-tensile steels, which can transmit tensile load better than casting materials.

The crankcase is usually drawn down at the sides to deep below the crankshaft bearings in order to provide sufficient rigidity. But sometimes a separate frame ("bed plate") may also be used which is designed to reduce weight and increase rigidity. Sometimes the side walls are additionally bolted horizontally at the sides

in the crankshaft-bearing area in order to achieve even greater rigidity. Bolted underneath is the oil pan, which can sometimes have a similarly bracing effect. However, this only applies to cast oil pans.

The piston bearing surface can be designed as a separate component (liner), particularly when replacement is not uncommon during the crankcase's service life, for example in commercial vehicles with high mileage ("wet liner" because coolant flows around it). However, the piston bearing surface can also be cast into the crankcase – particularly in cast-aluminum designs – or directly as a bearing surface in gray cast iron with strip hardeners.

Today's crankcases have attained a very high level of function integration and are therefore highly complex components. Frequently the oil-pump housing is also integrated in the casting, as are oil-cooler functions, the coolant-thermostat housing, the balancer-shaft locators, and all the brackets for the accessories.

Aside from the crankshaft, the crankcase is the component which contributes most to the engine weight. Many efforts are therefore made to reduce the weight by using materials other than the customary gray cast iron. These include aluminum materials, which, because of their thermally limited service life, are primarily used in smaller engines, particularly in passenger-car applications. Die casting has become an accepted cost-saving process for manufacturing aluminum, while sand casting is otherwise the regular process used. An "open-deck" construction is used to manufacture diecast crankcases efficiently, i.e. the joint face with the cylinder head features large pass-through cross-sections for the water jacket and the oil return. When cast-iron materials are used, on the other hand, the upper crankcase is closed to make the structure as rigid as possible.

Efforts are also being made to manufacture crankcases from magnesium in order to further reduce weight – but this comes at the expense of a high manufacturing cost. For this reason, a "hybrid" material concept is used where the bearing functions are performed by aluminum and the shielding functions by magnesium inserts. Further increases in strength are provided by other cast-iron variants, compacted graphite iron (CGI) being one such variant which has been more widely used in recent years. Special heat treatment and specific cooling can be used to create in cast iron high-strength (and hard-to-machine) spheroidal graphite instead of the customary lamellar graphite. This makes it possible to produce thinner walls and higher strength with reduced weight.

The crankcase is assigned the most important function of noise emission. Because the crankcase is often joined directly to the combustion chamber, absorbs the high periodic gas and inertial forces and has a relatively large surface, it is extremely important for this surface to be rigid. The standard design methods used to provide high rigidity are ribbing and barreling, which also help to prevent a membrane effect.

Variable compression ratio

A variable compression ratio is desirable both in gasoline engines to reduce the knock tendency close to full load and in diesel engines to reduce the peak pressure at full load. However, up to now no mechanical or hydraulic system has successfully managed to provide this variability in diesel engines. The reason for this is the high loads to which diesel engines are subjected on account of the higher peak pressures and higher pressure gradients compared with gasoline engines.

The designs for variable compression ratio (VCR) can be categorized as follows:

- Solutions which are integrated into the piston (hydraulic actuation of a piston upper section that moves axially in relation to a fixed lower section)
- Solutions which alter the length of the connecting rod (sliding solution, articulated connecting rod, eccentric connecting-rod bearing in the small or big connecting-rod eye)
- Formulations for an altered crankshaft position relative to the cylinder head (swiveling or axial movement of the crankcase, eccentric crankshaft bearing)
- Coupling lever for varying the effective piston stroke (multilink, crankshaft-drive lever system, displacement alteration with ε variation)

– Additional volumes which are increased or decreased (variable additional volume in the cylinder head)

Formulations for variable compression ratios in gasoline engines cover a range of approx. $\varepsilon = 7$ to 14.

Cylinder head

The main functions of the cylinder head are to supply the cylinder with fresh mixture, to discharge the exhaust gas, and to absorb the cylinder pressure forces.

Cylinder head

The cylinder head seals off the upper end of the crankcase and the cylinder barrel. It houses the gas-exchange valves and the spark plugs in gasoline engines, the fuel injectors in direct-injection gasoline engines, and the fuel injectors and possibly the sheathed-element glow plugs in diesel engines. Together with the piston, it also provides the desired combustion-chamber shape. In the vast majority of passenger-car engines, the entire valve gear is also mounted in the cylinder head.

The cylinder head performs a large number of functions and exhibits complex load profiles. The challenge faced here, as well as precisely controlling gas exchange and charge movement, which directly influence engine emissions, efficiency and power output, is above all to cool the hot side of the cylinder head and maintain control over the complex mechanical and thermal loads involved.

Based on the gas-exchange concepts, a distinction is made between two basic design configurations. In the counterflow cylinder head intake and exhaust passages open into the same side of the cylinder head (Figure 7b). This limits the space available for the intake and exhaust-gas passages, but due to the short flow tracts, this represents a substantial advantage in supercharged applications without intercooling. This design, with the gas supply and discharge tracts on a single side, also provides practical advantages in transverse-mounted engines.

In the crossflow cylinder head intake and exhaust passages are located on opposite sides of the engine (Figure 7a), providing a diagonal flow pattern for the intake and exhaust gases. This layout's advantages include more freedom in intake and exhaust-tract design as well as less complicated sealing arrangements. Typically, vehicle-specific boundary conditions are crucial to determining whether a counterflow or crossflow cylinder head is used.

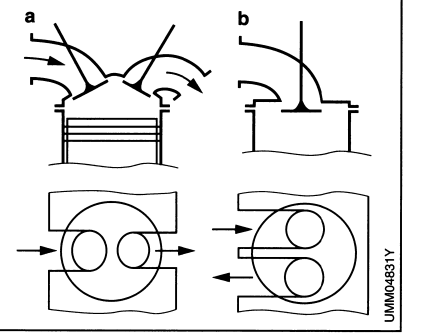
This decision also determines the valve arrangement used. Counterflow applications require a twisted valve arrangement in today's typical multi-valve designs. The main advantage of a twisted valve arrangement is that it is a simple way of delivering high cylinder-whirl values. A parallel valve arrangement requires greater efforts at the passage end to deliver a high cylinder whirl (Figure 8).

Today's passenger-car applications use almost exclusively a continuous single-piece cylinder head. In V-engines one cylinder head per cylinder bank is used. Individual cylinder heads are frequently still used in commercial-vehicle engines, which sometimes still feature a block-mounted camshaft with a push-rod assembly.

In contrast to passenger-car applications, in which the cylinder head is often made from aluminum, which as well as weight advantages improves heat transfer from the combustion chamber, commercial-vehicle applications utilize gray-cast-iron designs in view of the fact that

Figure 7: Cylinder-head identification according to intake and exhaust-tract location

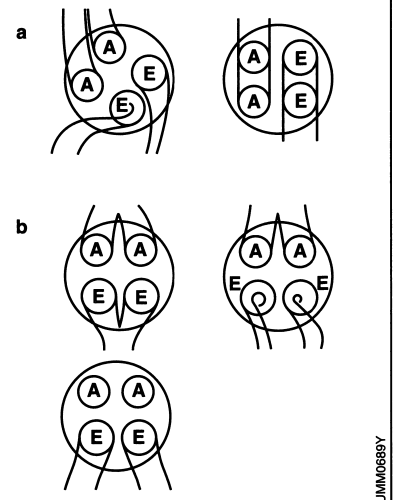
a) Crossflow design,
b) Counterflow design.



resistance to fatigue cannot be achieved by light alloys at hight peak pressures and with large cylinder diameters.

The cylinder head is subject to very high loads. On the one hand, frequently varying temperatures and the accompanying expansions give rise to stresses which can only be absorbed by appropriate design measures (individual cylinder head, suitable cavities in a multi-cylinder head). Thermomechanical strength is determined by the number of cold-hot load profiles.

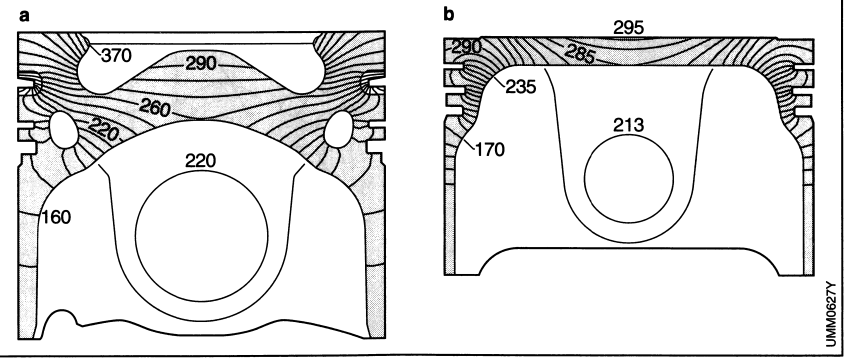
Figure 8: Valve arrangements
a) Twisted valve arrangements,
b) Parallel valve arrangements.
E Intake, A Exhaust.



In multi-valve designs the land between the exhaust valves is subjected to particularly high temperatures. At full load slightly critical temperatures above 350 °C are reached here. Similarly high temperatures are encountered in the recess-edge area (Figure 9). Gray-cast-iron designs are especially prone to the risk of hot-gas corrosion, an oxidation of the surface which can culminate in component-jeopardizing cracking.

The cylinder-head gasket, which is located between the crankcase and the cylinder head, performs an important function. Firstly, it must seal the combustion chamber even in the case of high cylinder pressures and, secondly, it must ensure the flow of engine oil and cooling water between crankcase and cylinder head. Typically, metal materials are used to seal the combustion chamber; these metals adapt to the surface by means of plastic deformation and establish the seal through their elasticity. Elastomer gaskets or geometrically optimized metal-sheet steel gaskets are used.

Figure 9: Piston operating temperatures in motor-vehicle engines at full load (schematic), values in °C
a) Passenger-car diesel-engine piston, 16 MPa ignition pressure, 58 kW/*l*,
b) Passenger-car gasoline-engine piston, 7.3 MPa ignition pressure, 53 kW/*l*.



To provide a uniform downward pressure on the cylinder-head gasket, it is possible to increase the number of cylinder-head bolts to such an extent that already commercial-vehicle engines with eight bolts per cylinder head are in use.

Engines without cylinder-head gaskets are the exception and require the crankcase deck and the underside of the cylinder head to be finished to very high levels of accuracy.

Valve gear

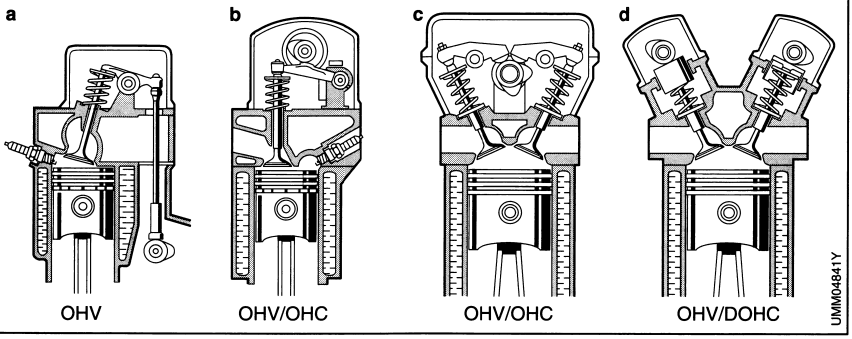
It is the function of the valve-gear assembly in a 4-stroke engine to permit and to control the exchange of gases in the internal-combustion engine. The valve gear includes the intake and exhaust valves, the springs which close them, the camshaft drive assembly, and the various force-transfer devices (Figure 10). In various widely-used designs, the camshaft is located in the cylinder head.

When the camshaft is installed within the block, the rocker arm is actuated not directly by the cam lobe, but by the intermediate action of a push rod and a lifter (push-rod assembly, Figure 10a).

In the case of the cam-follower or single rocker-arm assembly actuated by an overhead cam, the cam lobe's lateral and linear forces are absorbed and relayed by a cylinder-head mounted lever rocking back and forth between the cam lobe and the valve. In addition to transferring forces and absorbing lateral forces, the intermediate rocker arm can also be designed to magnify the cam pitch effect (Figure 10b).

In the case of the twin rocker-arm assembly actuated by an overhead cam, the rocker arm acts as the force-transfer element with its tilt axis located between the camshaft and the valve. Here, too, the rocker arm is usually designed as a cam-

Figure 10: Valve-timing designs
(Source: [1]).
a) Push-rod assembly, b) Single rocker-arm assembly,
c) Twin rocker-arm assembly, d) Overhead bucket-tappet assembly.
OHV Overhead valves,
OHC Overhead camshaft,
DOHC Double overhead camshaft.



lift multiplier to produce the desired valve travel (Figure 10c).

In the case of the overhead bucket-tappet assembly, a "bucket" moving back and forth in the cylinder head absorbs the cam lobe's lateral force, while transferring its linear actuating pressure to the valve stem (Figure 10d).

Valve arrangements

The valve-control arrangement and the design of the combustion chamber are closely interrelated. Today, nearly all valve assemblies are overhead units mounted in the cylinder head. In diesels and simpler spark-ignition (gasoline) engines, the valves are parallel to the cylinder axis, and are usually actuated by twin rocker arms, bucket tappets or single rocker arms. With increasing frequency, current spark-ignition (SI) engines designed for higher specific outputs tend to feature intake and exhaust valves which are inclined towards each other. This configuration allows larger valve diameters for a given cylinder bore, while also providing greater freedom for optimizing intake and exhaust passage design. Twin rocker-arm assemblies actuated by overhead cams are used most often here. High-performance and racing engines are increasingly using four valves per cylinder and overhead bucket-tappet valve assemblies.

The valve-timing diagram (Figure 11) for an engine shows the opening and closing times of the valves, the valve-lift curve, the maximum valve lift, and the valve velocity and acceleration.

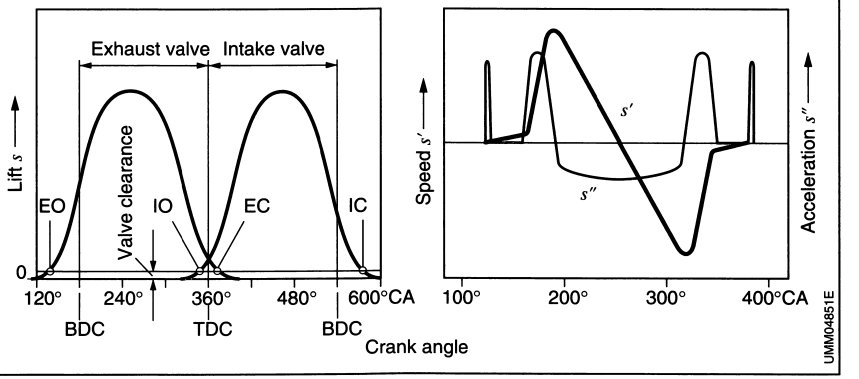
Typical valve-acceleration rates for passenger-car OHC (overhead camshaft) valve assemblies:
$s'' = 60...65$ mm $(b/\omega^2) \rightarrow 6{,}400$ m/s^2 at 6,000 rpm for single and twin rocker-arm assemblies.
$s'' = 70...80$ mm $(b/\omega^2) \rightarrow 7{,}900$ m/s^2 at 6,000 rpm for overhead bucket-tappet assemblies.
For heavy-duty commercial-vehicle engines with block-mounted camshafts:
$s'' = 100...120$ mm $(b/\omega^2) \rightarrow 2{,}000$ m/s^2 at 2,400 rpm.

Valve, valve guide and valve seat

The materials employed in manufacturing valves are heat and scale-resistant. The valve seat's contact surface is frequently hardened. Another proven method for improving the thermal-conductivity characteristics of exhaust valves is to fill their stems with sodium. To extend service life and improve sealing, valve-rotating systems (rotocaps) are also in common use.

The valve guides in high-performance engines must feature high thermal con-

Figure 11: Valve-timing diagram
s Valve lift, s' Valve velocity, s'' Valve acceleration.
IC Intake closes, IO Intake opens, EC Exhaust closes, EO Exhaust opens,
BDC Bottom dead center, TDC Top dead center.



ductivity and good antifriction properties. They are usually pressed into the cylinder head and are often supplemented by valve-stem seals at their cold ends for reducing oil consumption.

Valve-seat wear is generally reduced by making the valve-seat inserts of cast or sintered materials and shrink-fitting them into the cylinder head.

Cam-lobe design and timing dynamics

The cam lobe must be able to open (and close) the valve as far, as fast and as smoothly as possible. The closing force for the valves is applied by the valve springs, which are also responsible for maintaining contact between the cam lobe and the valve. Dynamic forces impose limits on cam and valve lift.

Rarely encountered are designs which use their own mechanism (Desmotronic) to restore the valves with the advantage of increased valve velocities and thus lift curves which more closely resemble the ideal rectangular case. Conventional springs must always ensure a frictional connection between camshaft, intermediate elements and valve while adhering to the permissible surface load. A slight clearance can be encountered only when cam lift is not active. Modern engines operate here with hydraulic valve-clearance compensation. Otherwise valve clearance must be regularly checked; measurements of 0.1 to 0.2 mm are common at the intake and exhaust valves.

Camshaft drive

To drive the valve gear, the camshaft must be connected to the crankshaft in such a way that one crankshaft revolution results in a one half camshaft revolution in a 4-stroke engine.

For passenger-car applications with overhead camshaft(s) either timing-chain or toothed-belt drives are used in modern engines. Both drives require a tensioning device, which acts on the free running length, the loose side, to prevent uncontrolled vibration.

In view of the fully utilized minimum space between piston and valve in modern engines, it is very important for the camshaft drive to operate safely and reliably, because in the event of damage contact can have serious consequences for the valve gear.

For this reason, the pressure pins for the tensioning rails in a chain drive or the toothed belt are wearing parts which must be replaced on a regular basis. Modern duplex chains are highly durable and are only subject to minimal wear. The drawbacks of a timing chain are its need to be lubricated and its elongation during operation. Some major advantages, however, are that it cannot jump off the gear wheel and it is not subject to the risk of cracking.

Transmission by gear wheels in a geartrain is typically used in commercial-vehicle engines or engines with block-mounted camshafts. In engines with overhead camshafts in particular, this is a very expensive solution, but one which provides for extending engine running times on account of the reliable transmission.

Expensive vertical shafts, whose axes run at right angles to the crankshaft and camshaft, are rarely used.

Oil supply

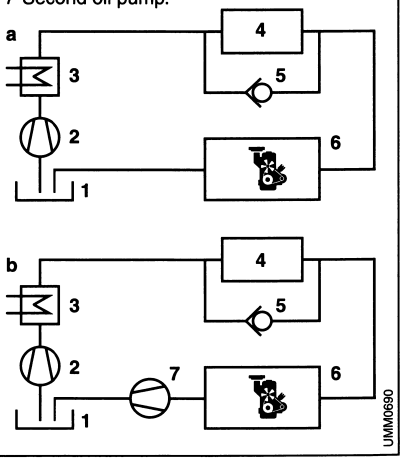
As well as lubricating the tribologically critical pairings of the crankshaft drive, the cylinder head and other components, the oil supply serves to remove local contaminants, combustion residues and wear particles which are filtered out in the oil-filter unit. Additional functions include the dissipation of heat in areas subject to thermal load, such as the friction bearings in the crankshaft drive or for oil-cooled pistons, and damping of vibrations in bearings.

With the system of forced-feed lubrication conventionally used (see also "Engine lubrication"), the oil pump, a gear pump that typically operates as a positive-displacement pump, delivers a defined volumetric flow from the oil sump through an oil-filter unit (Figure 12a). For safety reasons, the oil-filter unit is often equipped with a bypass valve and a pressure valve in a full-flow design. The oil cooler provided in engines subject to higher loads is cooled by either air or coolant. The engine oil flows though oil ducts and by the force of gravity back into the oil sump and into the oil pan typically located underneath the crankcase. That is why forced-feed lubrication is also known as wet-sump lubrication. In addition to the direct supply of oil by the oil pump, the rotational motion of the crankshaft effects a fine conditioning of the oil mist in the crankcase.

The opposite system to conventional wet-sump lubrication is the more expensive dry-sump lubrication, because it requires a second oil pump to remove the oil from the engine compartment (Figure 12b). The advantages are a constantly guaranteed lube-oil supply under conditions of high lateral acceleration or inclination and freedom in deciding on where to locate the oil-supply system. These give rise to lower engine heights and a possible increase in the engine-oil quantity, which is beneficial to engine cooling.

The mixture and total-loss lubrication frequently used in 2-stroke and Wankel engines is no longer relevant to present-day automobile applications.

Figure 12: Oil supply
a) Forced-feed lubrication (wet-sump lubrication),
b) Dry-sump lubrication.
1 Oil reservoir, 2 Oil pump, 3 Oil cooler, 4 Filter unit, 5 Bypass valve, 6 Engine with lubrication points, 7 Second oil pump.



Cooling

In order to avoid thermal overload, combustion of the lubricating oil on the piston's sliding surface, and uncontrolled combustion due to excessive component temperatures, the components surrounding the hot combustion chamber (cylinder liner, cylinder head, valves and in some cases the pistons themselves) must be intensively cooled (see also "Engine cooling").

As water has a high specific heat capacity and provides efficient thermal transition between the materials, most contemporary vehicle engines are water-cooled. Moreover, it is possible, under very hot local boundary conditions when the cooling medium is locally evaporated, to achieve through the heat of evaporation an additional cooling effect which together with the condensation behavior in colder adjacent areas reduces unwanted high local temperature gradients.

Air/water recirculation cooling is the most prevalent system (Figure 13). It comprises a closed circuit allowing the use of anticorrosion and antifreeze additives. The coolant is pumped through the engine and through an air/water radiator. The cooling air flows through the radiator in response to vehicle movement and/or

is forced through it by a fan. The coolant temperature is regulated by a thermostatic valve which bypasses the radiator as required.

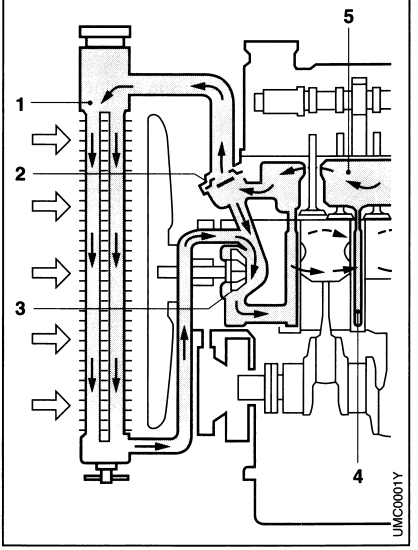
Air-cooled concepts still play a subordinate role today. Modern emission regulations call for effective engine cooling with simultaneously high specific power output. This can practically no longer be achieved with air cooling, a system which is indeed robust and low-maintenance, but is also not as effective. A crucial drawback of air-cooled concepts is their acoustic performance. The cooling fins act like a resonator and therefore result in a marked increase in sound emissions.

Crankshaft vibration damper

The crankshaft is subjected to high-frequency vibrations (2nd-order torsional vibrations), particularly in the region of resonant frequencies. Vibration dampers are therefore used to moderate dangerous crankshaft resonance phenomena.

Figure 13: Water cooling system with coolant circuit

1 Radiator,
2 Thermostat,
3 Water pump,
4 Water passages in cylinder block,
5 Coolant passages in cylinder head.



The choice of vibration-damper design ranges from simple flywheel masses, which through elastically damping layers – e.g. rubber – provide a countervibration system, through more sophisticated oil-damped flywheel masses, in which the viscosity of the oil and friction surfaces provide the damping effect, up to complex pendulum adsorbers. The vibration damper is generally mounted at the free crankshaft end (front).

1st-order torsional vibrations, i.e. vibrations resulting from ignition and inertial forces, which can result in excitation of transmission and input-shaft vibrations, are reduced by dual-mass flywheels or other damping technologies. Phenomena such as clutch grabbing, transmission chatter and starting jerks can be optimized by such vibration absorbers, since all the following components – clutch, transmission and input shafts – likewise represent an oscillatory system. In a dual-mass flywheel the two flywheel masses are coupled to each other by means of springs. Torsional-vibration dampers are also commonly featured in clutches – both components must naturally be tuned together.

Design and tuning are difficult to simulate and must in the final analysis be conducted in tests, both in the vehicle with the entire drivetrain (1st order: low-frequency drivetrain resonance phenomena) and on the test bench (2nd order: high-frequency crankshaft resonances).

Reciprocating-piston engine types

The internal-combustion engine is characterized by its great variability and application-specific adaptability. In particular, the arrangement of the individual cylinders provides for numerous variants.

Arrangements

In principle, a huge number of possible cylinder arrangements is conceivable. Few of these arrangements have proven to be particularly effective for automotive applications (Figure 14).

Radial engines are not suitable for automotive applications on account of their height and are therefore not used.

It is important to note the difference between the opposed-cylinder (boxer) engine and the V-engine with 180° included angle between cylinder banks. In the boxer engine the pistons always move in opposite directions, which is why the inertial forces cancel each other out. In the V-engine, on the other hand, the pistons of two opposed cylinders always move in the same direction.

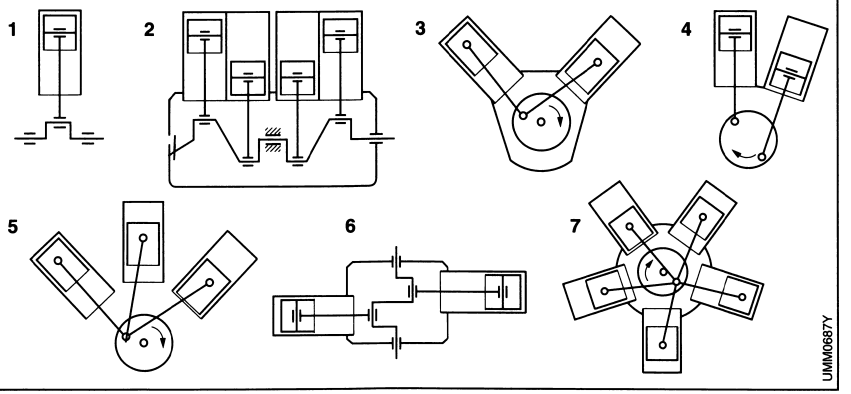
It is also important to note the design of V- and W-engines, in which the connecting-rod/crankshaft link (large connecting-rod eye) of associated cylinders is effected at slightly offset crank pins.

In multi-piston units compression is typically generated by several working pistons. In the U-engine the pistons undertake an approximately parallel motion. In the opposed-piston engine the pistons undertake a typically opposed motion. Engine variants with more than one piston per combustion chamber have not gained acceptance. Weight and engine size are not suitable for automotive applications.

Definitions

The following definitions in accordance with DIN 73021 [2] as viewed looking at the end opposite the power-output end apply only to motor-vehicle engines. For internal-combustion engines for general and marine use, the reverse direction – i.e. as viewed looking at the power-output end – is standardized (ISO 1204 [3]).

Figure 14: Reciprocating-piston engine types
1 Single cylinder (individual cylinder),
2 In-line engine (arrangement of individual cylinders in line),
3 V-engine (arrangement of individual cylinders in two banks in a V-shape),
4 VR-engine (arrangement of individual cylinders in two banks in a V-shape, where the two banks are particularly close to one another),
5 W-engine (arrangement of individual cylinders in three banks in a W-shape),
6 Opposed-cylinder (boxer) engine (arrangement of individual banks opposing one another),
7 Radial engine (radial arrangement in one or more planes).



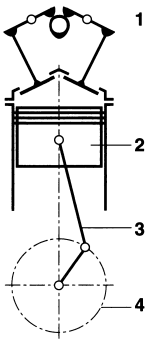
Clockwise rotation
The direction of rotation is clockwise, as viewed looking at the end opposite the power-output end.

Counterclockwise rotation
The direction of rotation is counterclockwise, as viewed looking at the end opposite the power-output end.

Numbering the cylinders
The cylinders are numbered consecutively 1, 2, 3, etc. in the order in which they would be intersected by an imaginary reference plane, as viewed looking at the end of the engine opposite the power-output end. This plane is located horizontally to the left when numbering begins; the numerical assignments then proceed clockwise about the longitudinal axis of the engine. If there is more than one cylinder in a reference plane, the cylinder nearest the observer is assigned the number 1, with consecutive numbers being assigned to the following cylinders. Cylinder 1 is to be identified by the number 1.

Firing sequence
The firing sequence is the sequence in which combustion is initiated in the cylinders. Engine design configuration, uniformity in ignition intervals, ease of crankshaft manufacture, optimal crankshaft load patterns, etc. all play a role in defining the firing sequence.

Figure 15: Crankshaft assembly for reciprocating-piston engine (principle)
1 Valve gear,
2 Piston,
3 Connecting rod,
4 Crankshaft.



Crankshaft-assembly design

Crankshaft-drive kinematics
The kinematics of the single-cylinder crankshaft assembly can be determined from the geometrical assignments of piston and piston-pin axis, connecting rod and crankshaft (crankshaft radius equal to one half stroke) (Figure 15). If piston travel x at top dead center is assumed to be zero, the following is obtained with crank radius r and connecting-rod length l (Figure 16):

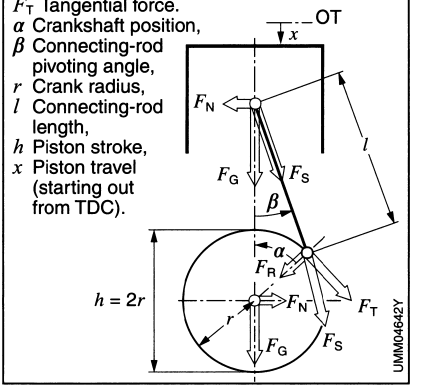
$$x = r\,(1 - \cos\alpha) + l\,(1 - \cos\beta)\ .$$

Where
$r \sin\alpha = l \sin\beta$
and
$\lambda = r/l$
the following is obtained:

$$x = r\left(1 - \cos\alpha + \frac{1}{\lambda}\left(1 - \sqrt{1 - \lambda^2 \sin^2\alpha}\right)\right)$$

An offset of the piston pin is performed by some manufacturers. Noise and friction advantages can be expected on account of the changing piston location and depending on the position of the connecting rod. The offset can be effected either in the piston by shifting the piston pin from the center position or by offsetting the crankshaft.

Figure 16: Gas-force components shown on a basic crankshaft drive
F_G Piston-pin force (gas force),
F_S Connecting-rod force, F_R Radial force,
F_N Piston normal force,
F_T Tangential force.
α Crankshaft position,
β Connecting-rod pivoting angle,
r Crank radius,
l Connecting-rod length,
h Piston stroke,
x Piston travel (starting out from TDC).



If the offset for positive crank angles is assumed to be positive and the quantity

$$\delta = \frac{\text{Offset}}{\text{Connecting-rod length}}$$

is introduced, this leads to the following function of the piston travel:

$$x = r\left(1 - \cos\alpha + \frac{1}{\lambda}\left(1 - \sqrt{1 - (\lambda \sin\alpha - \delta)^2}\right)\right)$$

Figure 17 shows by way of example the influence of the stroke/connecting-rod ratio and the offset. However, the differences with normal offset values in the millimeter range ($\delta < 0.04$) are clearly smaller.

By developing the root function into a Taylor series (around $x = 0$: Maclaurin series) and replacing the powers of the trigonometric functions with the multi-harmonic functions, the following expression can be obtained [4]:

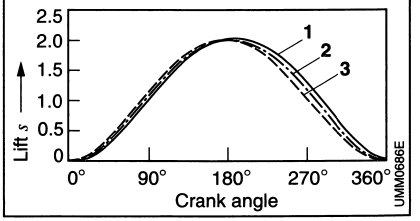
$$x = r\,[1 + \tfrac{1}{4}\lambda + \tfrac{3}{64}\lambda^3 + \ldots - \cos\alpha - (\tfrac{1}{4}\lambda + \tfrac{3}{64}\lambda^3 + \ldots)\cos 2\alpha + (\tfrac{3}{64}\lambda^3 + \ldots)\cos 4\alpha + \ldots].$$

This expression shows that there are higher harmonics on account of the crankshaft-drive kinematics, and these are also called engine orders (multiples of the engine speed).

Because normal values for λ are around 0.3, the higher-level λ terms are readily ignored and the following simplified function is assumed for further calculations:

$$x = r\,[1 + \tfrac{1}{4}\lambda - \cos\alpha - \tfrac{1}{4}\lambda \cos 2\alpha\,].$$

Figure 17: Piston-travel function
1 $\lambda = \infty$,
2 $\lambda = 0.3$,
3 $\lambda = 0.3$, $\delta = 0.1$.



This simplification may not be used if detailed vibration analyses and resonance phenomena are to be examined.

From the simplified equation the following piston speeds v and accelerations a are obtained, where the angular velocity $d\alpha/dt = \omega = 2\pi n$ (n speed) was introduced:

$$v = r\omega\,(\sin\alpha + \tfrac{\lambda}{2}\sin 2\alpha),$$

$$a = r\omega^2\,(\cos\alpha + \lambda\cos 2\alpha).$$

Here too there are higher harmonics (orders), which must not be disregarded when resonances are examined.

Crankshaft-assembly kinetics

The forces acting on the crankshaft assembly and the resulting moments can initially be derived without inertial forces as follows (Figure 16).

The piston-pin force results as gas force from the combustion-chamber pressure and the piston area. The following applies:

$$F_G = (p - p_{KGH})\,A_{piston}.$$

The connecting-rod force is obtained from the vectorial analysis of the piston-pin force in the direction of the connecting rod. The following applies:

$$F_S = \frac{F_G}{\cos\beta} = \frac{F_G}{\sqrt{1 - \lambda^2 \sin^2\alpha}}$$

The piston normal force F_N is the vectorial component of the piston-pin force normal to the cylinder wall and for balancing the connecting-rod force:

$$F_N = F_G \tan\beta = \frac{F_G\,\lambda \sin\alpha}{\sqrt{1 - \lambda^2 \sin^2\alpha}}.$$

This force contributes significantly to the friction of the piston on the cylinder barrel. The side which is touched by the piston after top dead center on account of the combustion pressure is termed the major thrust face, the opposite side the minor thrust face. The greatest friction is therefore created shortly after TDC on the major thrust face.

The tangential force at the crankshaft crank pin contributes to an acceleration of the crankshaft and thus to the build-up of torque at the crankshaft. It is created from the vectorial analysis of the connecting-rod force.

$$F_T = \frac{F_G \sin(\alpha+\beta)}{\cos\beta} = F_G\left(\sin\alpha + \frac{\lambda}{2}\sin 2\alpha/\sqrt{1-\lambda^2\sin^2\alpha}\right).$$

Here too the radicand can be simplified in a series expansion as

$$F_T \approx F_G\,(\sin\alpha + \frac{\lambda}{2}\sin 2\alpha).$$

The radial force F_R at the crankshaft crank pin is:

$$F_R = F_G\frac{\cos(\alpha+\beta)}{\cos\beta} = F_G\left(\cos\alpha + \lambda\frac{\sin^2\alpha}{\sqrt{1-\lambda^2\sin^2\alpha}}\right)$$

or approximately

$$F_R \approx F_G\,(\cos\alpha - \frac{\lambda}{2} + \frac{\lambda}{2}\cos 2\alpha).$$

The inertial forces can be split into oscillating and rotating components. The piston, piston-ring and piston-pin weights m_K belong to the oscillating component and can be punctually assigned to the piston pin.

The crankshaft web with crank pin belong to the rotating component. Here the mass is usually reduced to the crank radius and punctually assigned to the crank-pin center point. The following applies:

$$m_W = \Sigma\frac{m_i\, r_{si}}{r},$$

where m_i is the respective mass component of web, pin, etc. and r_{si} is the corresponding center-of-mass radius.

Because of the connecting rod's swinging motion, it is useful to split the connecting-rod mass into an oscillating component and a rotating component. This is done exactly where the connecting-rod-mass center of gravity and mass moment of inertia are known by assuming two dynamically equal individual masses at the small and big connecting-rod eyes and by calculating the equilibrium formulation of force, moment and mass inertia. Usually one third of the connecting-rod mass m_{Pl} can be assumed to be oscillating and two thirds to be rotating. Then, with $m_o = m_K + 1/3\ m_{Pl}$ as the oscillating mass and the piston acceleration (see below), an oscillating inertial force is obtained at:

$$F_o \approx m_o\, r\, \omega^2\,(\cos\alpha + \lambda\cos 2\alpha).$$

The oscillating inertial force therefore increases quadratically with the engine speed ($\omega = 2\pi n$) and has a first-order component and a lower second-order component.

The rotating inertial force is obtained as centrifugal force from the reduced mass $m_r = m_W + 2/3\ m_{Pl}$ and the rotation speed at:

$$F_r = m_r\, r\, \omega^2.$$

The rotating inertial force likewise increases with the square of the speed, but has not higher orders. The rotating inertial force can therefore be easily balanced by counterweights which rotate at engine speed. The rotational irregularities of the crankshaft are so small compared with these forces that they can be ignored in the balancing of masses.

As shown in the crankshaft-assembly kinematics, higher harmonics – higher engine orders – occur on account of the crankshaft-drive geometry. Aside from the 1st and 2nd engine orders, however, the amount of the 4th order and the higher components decreases rapidly, and is generally ignored for a balancing of masses.

Balancing of masses in a single-cylinder engine

The component of rotating masses in a single-cylinder engine can be fully balanced by appropriate counterweights on the crank pin. Weights are normally provided for on both sides and the weights must be balanced merely with the center-of-mass radius. The oscillating forces can be represented with force vectors (Figure 18) when they are modeled rotating in the opposite direction in each case with half the amount.

Therefore two contra-rotating shafts with weights can be used to balance the oscillating inertial forces. The horizontal component then cancels out and at least the first order of the oscillating inertial-force component can be compensated.

A virtually full balancing of masses requires further balancer shafts, which must rotate at twice the engine speed in order to balance completely the second-order component of the oscillating mass component.

A compromise is often reached due to the fact that the expenditure associated with contra-rotating shafts is high and at least with the first order already not inconsiderable weights would have to be realized in the shafts. For example, half the oscillating mass can be integrated in the counterweights. The free inertial forces acting outwards are now reduced in the cylinder longitudinal direction by one half; however, additional transverse forces are now created on account of the excessive rotating components (see Table 2). A half compensation of this kind is termed a 50% balancing rate. Usual figures are 100% rotatory and 50% oscillatory balancing of masses.

Balancing of masses in a multiple-cylinder engine

The inertial forces of a multiple-cylinder engine are composed of the inertial forces of each individual cylinder which are superimposed according to the crankshaft throw. In addition, free moments of inertia are created on account of the cylinder spacing. Table 3 shows a summary of the possible transverse and longitudi-

Figure 18: Complete balancing of 1st- and 2nd-order masses in a single-cylinder assembly

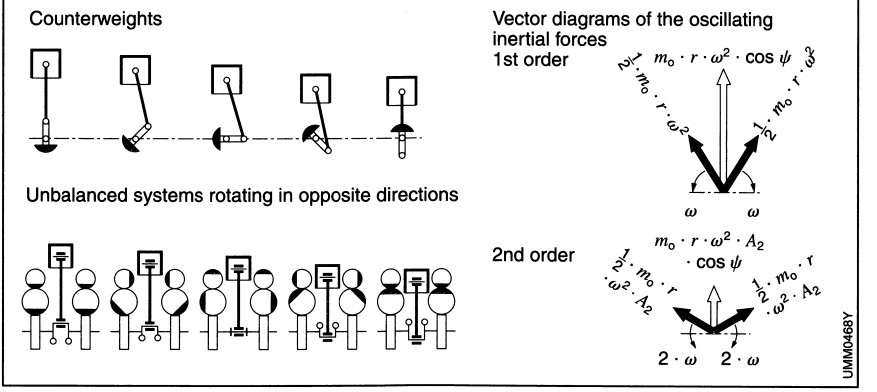
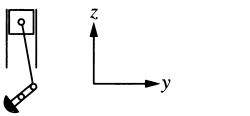

Counterweights
Unbalanced systems rotating in opposite directions
Vector diagrams of the oscillating inertial forces
1st order
$m_o \cdot r \cdot \omega^2 \cdot \cos \psi$
$\frac{1}{2} \cdot m_o \cdot r \cdot \omega^2$
$\frac{1}{2} \cdot m_o \cdot r \cdot \omega^2$
ω ω
2nd order
$m_o \cdot r \cdot \omega^2 \cdot A_2 \cdot \cos \psi$
$\frac{1}{2} \cdot m_o \cdot r \cdot \omega^2 \cdot A_2$
$\frac{1}{2} \cdot m_o \cdot r \cdot \omega^2 \cdot A_2$
$2 \cdot \omega$ $2 \cdot \omega$
UMM0468Y

Table 2: Balancing of masses on a single-cylinder engine, dependent on the balancing rate

 UMM0470Y		Balancing rate		
		0%	50%	100%
Size of counterweight	$m_G \triangleq$	m_r	$m_r + 0.5m_0$	$m_r + m_0$
Residual inertial force (z) 1st order	$F_{1z} =$	$m_0 \cdot r \cdot \omega^2$	$0.5 \cdot m_0 \cdot r \cdot \omega^2$	0
Residual inertial force (y) 1st order	$F_{1y} =$	0	$0.5 \cdot m_0 \cdot r \cdot \omega^2$	$m_0 \cdot r \cdot \omega^2$

nal tilting moments that occur and the free inertial forces.

The mutual balancing of the inertial forces is one of the essential factors determining the selection of the crankshaft's configuration, and with it the design of the engine itself. The inertial forces are balanced if the common center of gravity for all moving crankshaft-assembly components lies at the crankshaft's midpoint, i.e. if the crankshaft is symmetrical (as viewed from the front). This is represented with the 1st-order and 2nd-order star diagrams (Table 4).

The 2nd-order star diagram for the four-cylinder in-line engine is asymmetrical, meaning that this order is characterized by free inertial forces. These forces can be balanced using two countershafts rotating in opposite directions at double the rate of the crankshaft (Lanchester system).

Table 5 shows a summary of free forces and moments based on balancing of masses for different numbers of cylinders and throw variants.

Torsional force

Moving masses are exposed to constantly changing acceleration, thereby generating inertial forces. The cyclically occurring pressure forces in the cylinders are termed gas forces. Both form in total internal and external forces and moments. The internal forces and moments must be

Table 3: Transverse and longitudinal tilting moments, free inertial forces in multiple-cylinder engines

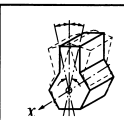
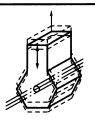
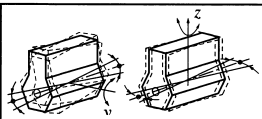
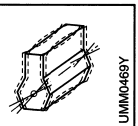
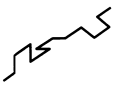
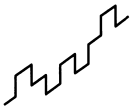
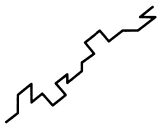
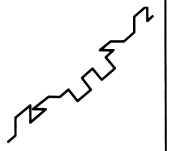
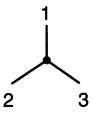
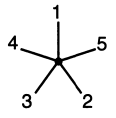
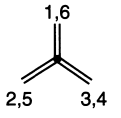
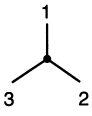
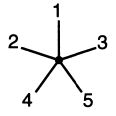
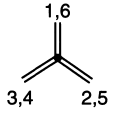
Forces and moments at the engine				 UMM0469Y
Designation	Oscillating torque, transverse tilting moment, reaction torque	Free inertial force	Free inertial moment, longitudinal tilting moment about the y-axis (transverse axis) ("pitching" moment) about the z-axis (vertical axis) ("rolling" moment)	Internal bending moment
Cause	Tangential gas forces as well as tangential inertial forces for the ordinals 1, 2, 3 and 4	Unbalanced oscillating inertial forces 1st order in 1- and 2-cylinders; 2nd order in 1-, 2- ,4-cylinders	Unbalanced oscillating inertial forces as a composite of 1st- and 2nd-order forces	Rotating and oscillating inertial forces
Design factors	Number of cylinders, ignition intervals, displacement, p_i, ε, p_z, m_0, r, ω, λ	Number of cylinders, crank configuration m_0, r, ω, λ	Number of cylinders, crank configuration, cylinder spacing, counterweight size influences inertial-torque components about the y- and z-axes m_0, r, ω, λ, a	Number of throws, crank configuration, engine length, engine-block rigidity
Remedy	Can only be compensated for in exceptional cases	Free mass effects can be eliminated by rotating balancing systems, however, this process is complex and therefore rare; crank sequences with limited or no free mass effects are preferable		Counterweights, rigid engine block
	Shielding of the environment through flexible engine mounts (in particular for orders ≥ 2)			

Table 4: Star diagram for in-line engines

	3-cylinder	4-cylinder	5-cylinder	6-cylinder
Crank sequence				
Star diagram 1st order				
Star diagram 2nd order				

absorbed by the components, especially the crankshaft and crankcase, while the external effects act via the engine bearings on the support structure and impart oscillations to the chassis or engine foundation.

If the periodic gas force acting on the piston and the periodic mass inertial forces acting on the piston, connecting rod and crankshaft assembly are grouped together, they generate a sum of tangential force components at the crankshaft journal. When multiplied by the crank radius, this produces a periodically variable torque value.

With multiple-cylinder engines, the tangential-pressure curves for the individual cylinders are superimposed with a phase shift dependent on the number of cylinders, crankshaft-journal configuration, crankshaft design, and the firing sequence. The resulting composite curve is characteristic for the engine design, and covers a full working cycle (i.e. two crankshaft rotations for 4-stroke engines, Figure 18). It can be depicted in a torsional-force diagram. This variable torsional force and the resulting torque generate, depending on the available moment of inertia J, a variable rotation speed ω:

$$\frac{d\omega}{dt} = \frac{M(t)}{J}$$

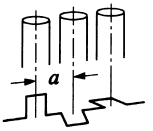
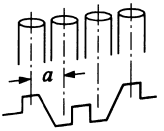
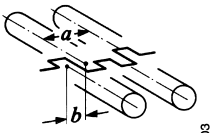
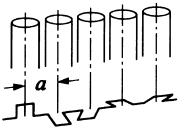
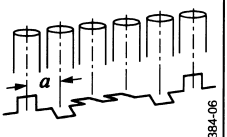
with all the superimposed and newly created engine orders (there are also half orders). This deviation from the speed constant is called the coefficient of cyclic variation and is defined as follows:

$$\delta_S = \frac{\omega_{max} - \omega_{min}}{\omega_{min}}.$$

This coefficient of cyclic variation is reduced to a level suitable for the application by energy-storage mechanisms, such as a flywheel mass for example. The torsional vibrations which can be traced back to the torsional force depicted are also called 1st-order torsional vibrations. These are not to be confused with the (high-frequency) torsional vibrations resulting from elastic deformations and natural resonances of the crankshaft (Table 5), which are referred to as 2nd order.

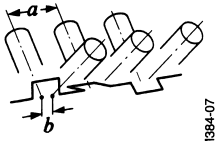
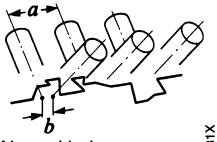
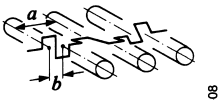
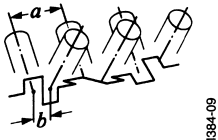
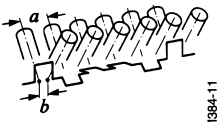
Table 5: Free forces and moments of the 1st and 2nd order, and ignition intervals of the most common engine designs

$F_r = m_r\, r\, \omega^2 \quad F_1 = m_0\, r\, \omega^2 \cos\alpha \quad F_2 = m_0\, r\, \omega^2\, \lambda \cos 2\alpha$

Cylinder arrangement	Free forces of 1st order[1]	Free forces of 2nd order	Free moments of 1st order[1]	Free moments of 2nd order	Ignition intervals
3-cylinder					
 VM384-01 In-line, 3 throws	0	0	$\sqrt{3} \cdot F_1 \cdot a$	$\sqrt{3} \cdot F_2 \cdot a$	240°/240°
4-cylinder					
 VM384-02 In-line, 4 throws	0	$4 \cdot F_2$	0	0	180°/180°
 VM384-03 Opposed-cylinder (boxer), 4 throws	0	0	0	$2 \cdot F_2 \cdot b$	180°/180°
5-cylinder					
 VM384-05 In-line, 5 throws	0	0	$0.449 \cdot F_1 \cdot a$	$4.98 \cdot F_2 \cdot a$	144°/144°
6-cylinder					
 VM384-06 In-line, 6 throws	0	0	0	0	120°/120°

[1] Without counterweights.

Table 5 (continued): Free forces and moments of the 1st and 2nd order, and ignition intervals of the most common engine designs

Cylinder arrangement	Free forces of 1st order [1]	Free forces of 2nd order	Free moments of 1st order [1]	Free moments of 2nd order	Ignition intervals
6-cylinder (continued)					
 V 90°, 3 throws	0	0	$\sqrt{3} \cdot F_1 \cdot a$ [2]	$\sqrt{6} \cdot F_2 \cdot a$	150°/90° 150°/90°
 Normal balance V 90°, 3 throws, 30° crank offset	0	0	$0.4483 \cdot F_1 \cdot a$	$(0.966 \pm 0.256) \cdot \sqrt{3} \cdot F_2 \cdot a$	120°/120°
 Opposed-cylinder (boxer), 6 throws	0	0	0	0	120°/120°
 V 60°, 6 throws	0	0	$3 \cdot F_1 \cdot a/2$	$3 \cdot F_2 \cdot a/2$	120°/120°
8-cylinder					
V 90°, 4 throws in two planes	0	0	$\sqrt{10} \cdot F_1 \cdot a$ [2]	0	90°/90°
12-cylinder					
 V 60°, 6 throws	0	0	0	0	60°/60°

[1] Without counterweights, [2] Can be completely balanced by using counterweights.

Tribology and friction

The piston with piston rings and bearing surface is a self-contained and highly complex tribological system. This also applies to the piston pin and the crankshaft friction bearings. Detailed simulation formulations are necessary in view of the periodically changing application of force both in terms of amount and direction.

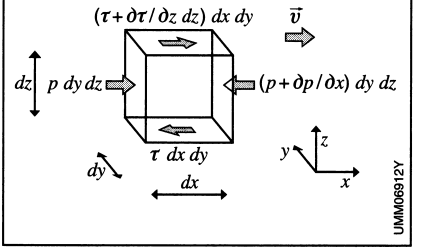
To be able to ensure wear resistance, it is necessary to build up a hydrodynamic lubricating film which is higher than the surface roughness of the two contact surfaces (see also Stribeck curve). The oil film either is already present in the lubricating gap (friction bearing with compressed-oil supply) or must be built up dynamically as an oil wedge (cambered piston ring on liner). However, the oil wedge always collapses when the relative speed of the contact surfaces assumes a value of zero. This is the case for piston rings at the piston reversal points (top and bottom dead centers), and for the piston pin between these positions. It is therefore also important to ensure that a sufficient oil-retaining volume can be maintained in the contact surfaces in order to retain by forces of adhesion in the short time of standstill as much oil as to avoid seizing or minimize increased wear. In the tribological system for the piston this is bearing-surface honing, and for the piston pin shaping of the piston-pin bearing.

Traced back to Newton is the formulation to the effect that the friction between two adjacent oil elements is virtually independent of the prevailing pressure and only proportional to the change in speed from one element to the next.

$$\tau = \frac{\eta\, dv}{dz}.$$

$\tau = F/A$ Shear stress,
F Shear force,
A Contact surface,
v Speed,
z Coordinate direction normal to speed direction v.

Figure 19: Pressure forces in a one-dimensional stress state



Sometimes the kinematic viscosity ν is also used; this corresponds to the dynamic viscosity referred to density and is defined by

$$\nu = \frac{\eta}{\rho}\ [\mathrm{m^2/s}].$$

The dynamic viscosity η [N s/m^2 or Pa · s] decreases sharply with temperature. It is important to see to it that the oil temperature ensures a viscosity η which facilitates sufficient shear stress to build up a supporting oil film. It is also necessary to provide for a sufficiently high relative speed v on the part of the contact surfaces. It can be recognized from the formula that hydrodynamic shear stress collapses when the speed assumes a value of zero (reversal point of the piston, pivoting-angle end of the small connecting-rod eye). If can further be recognized that the lubricating film Δy must not become too large, as otherwise the shear stress will become too small and the lubricating film will break.

The equilibrium of forces for the simplified case of a one-dimensional stress state in the x direction (Figure 19) shows that $\partial p/\partial x = \partial \tau/\partial z$ must apply and thus $\partial p/\partial x = \eta\, \partial^2 v/\partial z^2$. It is thus demonstrated on the one hand that the pressure loss of a flow is dependent on the viscosity and the speed. In addition, the speed profile can thus be derived.

On account of the adhesion condition which states that the speed of the oil on the wall is identical to the wall speed, a speed distribution (linear for laminar flow) and a pressure distribution build up in the lubricating gap (see Friction bearings).

It is attempted with the aid of (thermo-) elastohydrodynamic simulation (EHD, elastohydrodynamics) to determine the lubricating-film thickness by solving the mathematical equations in the multi-dimensional case for mass balance and momentum equilibrium.

- Mass balance: supplied oil mass and oil quantity escaping through the bearing gap in equilibrium.
- Momentum equilibrium: equilibrium of forces from direct stress (pressure force from oil pressure), shear stress (viscosity force from oil viscosity) including apparent viscosity stresses on account of turbulent flow and the inertial force (acceleration component).

The equation system is established as a Reynolds differential equation or a Navier-Stokes equation to solve the flow problem. By using a finite-element simulation, the bearing deformations are added and the highly complex equations (because they are simultaneous) are numerically solved.

An adequate supply of oil must be guaranteed in the bearings, since generally the oil pressure in the bearings (> 100 bar) far exceeds the static oil pressure provided by the oil pump (< 10 bar). This is achieved by means of pressure-oil bores and special grooves in friction bearings or by means of initial forces at the piston rings.

Attempts have also been made in the meantime to simulate the case of mixed friction. The influence of surface fine geometry (consideration of roughness on microhydrodynamics by "flow tensors") is taken into account in an extended Reynolds differential equation. By using contact-pressure models of rough surfaces, it is possible to determine, as well as the hydrodynamic contact-area ratios, the solid contact-area ratio and to take this into account accordingly. Intensive work is being conducted on models to determine the running-in behavior of contact surfaces. This will finally enable the wear behavior to be assessed.

Because the shear stresses and thus also the frictional forces are proportional to the speed, the share of friction of the piston assembly at the dead centers and that of the piston-pin bearing is small. Instead, the wear parameters are decisive here.

Sliding bearings

The bearings most commonly encountered in an internal-combustion engine are friction (sliding) bearings. Crankshaft main bearings, big-end bearings, piston-pin bearings, and camshaft bearings are generally oil-lubricated friction bearings (see Friction bearings).

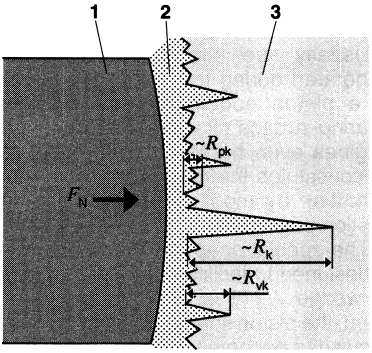
It is important to ensure that the seating pressure is precisely configured in both split bearings (crankshaft, connecting rod) and non-split bearings (piston pin, camshaft) so as to prevent them from jamming and rotating also at higher temperatures or under emergency-lubrication conditions. However, material overloading must also be reliably prevented.

The lubricating gap of the crankshaft main bearings changes in response to the constantly varying bearing forces (see also Crankshaft-assembly design). This generates a kind of pumping effect and the oil in the bearing is constantly replaced.

The swinging motion of the connecting rod effects an increased relative speed when the piston travels up the cylinder and a reduced relative speed when the piston travels down the cylinder in the lubricating gap of the big-end bearing or crank pin. This alone causes oil to be pumped into the oil gap. Furthermore, naturally periodic forces acting on the bearing cause the connecting rod to change position constantly and thereby assist the build-up of lubricating-film pressure.

The small-end bearing is merely subject to a swinging motion. A supply of oil is therefore particularly important, but difficult to achieve. In practical terms, mixed friction cannot be avoided. If the arrangement involves a floating pin bearing (bearing in the piston and connecting rod), the pin performs a slight rotary motion in the

Figure 20: Roughness values and oil-film theory (roughness definitions, see DIN 4760 [5])
1 Piston, 2 Oil film, 3 Liner.



piston bearing of just a few degrees – here too a good supply of oil to the bearings must be ensured.

Oil is generally supplied as splash oil, and partly also via a specific locating bore or via a pressure-oil supply from the crankshaft through grooves in the crankshaft-journal bearing and a vertical bore in the connecting rod.

Rolling pairs

Ball, roller or needle bearings are used in 2-stroke engines or sometimes in camshafts (see Roller bearings).

A roller follower or the roller of an injection-pump plunger-and-barrel assembly is a special case in point. The system of rotating roller and moving cam is a rolling-element pair with linear contact. A characteristic load variable is the Stribeck rolling-contact pressure:

$$K = \frac{F}{D_l \, l_{eff}},$$

F Force load,
$\varphi = D_l/D$ Osculation coefficient,
D_l Substitute diameter,
D Roller diameter,
l_{eff} Effective roller width.

The effective substitute diameter D_l is obtained from the instantaneous bend radius of the cam R_{cam} and the roller radius R_{roller}:

$$D_l = 2 \, \frac{R_{cam} \, R_{roller}}{R_{cam} + R_{roller}}.$$

It takes into account the osculation of the two rolling bodies.

The Hertzian stress p_{Hertz} is also relevant:

$$p_{Hertz} = \sqrt{\frac{KE}{2.86}}.$$

with modulus of elasticity E:

$$E = 2 \, \frac{E_{cam} \, E_{roller}}{E_{cam} + E_{roller}}.$$

By slightly angling the roller in relation to the cam or by offsetting the stroke function, it is possible to ensure that an oil wedge can be built up and the roller rotates continuously. Pitting is a typical indicator of overloading in rolling pairs that are subject to high loads. It results from molecular deformations of the metal and possible embedding of oil, which gives rise to extremely high local pressures and plastic deformations or material erosion. Harder materials and decreasing the load reduce the tendency to pitting.

Piston ring and bearing surface

The piston rings together with the bearing surface constitute a complex tribological system. On the one hand, the pressure force of the piston ring is very much dependent on the cylinder pressure and thus on the crank angle and the power stroke. On the other hand, the piston rings make axial secondary movements in their grooves which also have a significant impact on the lubricating-film thickness.

The oil is generally applied as splash oil to the bearing surface (crankshaft churning or splash oil for piston cooling and pin lubrication) and allowed back through the oil control ring as a thin film. The oil-retaining volume of the liner is dependent on the honing. This includes the surface roughness (furrow depth) and the honing pattern (cross-furrow angle).

Honing the bearing surface serves two functions. Firstly, it is intended to ensure with the aid of honing furrows of sufficient depth that oil can be retained by adhesion forces for an adequate length of time (Figure 20). Secondly, surface roughness must be so low as to prevent contact with the piston ring. Various roughness

values are used to characterize honing; these values are standardized on a generic level in DIN 4760 [5] and in detail in DIN EN ISO 13565 [6]. For initial wear and initial friction the R_{pk} value ("reduced peak height", see DIN 4760 and DIN EN ISO 11562 [7] or DIN EN ISO 13565 for an explanation) is decisive – but the value is swiftly and drastically minimized after run-in (e.g. from a few micrometers to 0.2 to 0.8 µm). The R_{vk} value (1.0 to 3.5 µm) is the decisive parameter for the possible oil-retaining volume after run-in. It is to be noted that the honing usually takes the form of a cross-shaped furrow system. In this form, the furrows constitute a communicating system of oilways and the oil pressure is partly forced out of the gap, but this also permits a uniform distribution of the oil film over the surface. Flat honing angles (referring to the angle formed by the cross-shaped furrows) reduce the amount of oil forced out, but there is a tendency to seizing on account of the unsatisfactory axial oil distribution. Steep honing angles result in increased oil consumption. Honing angles between 30° and 90° are customary.

Mixed friction occurs at any rate at the piston reversal points. The least that is required in this area is micro-pockets of oil, which can be produced for example by laser honing or by bringing out amounts of alloy during the honing procedure.

Usually, such high temperatures are generated during the combustion stroke of a piston engine that much of the bearing-surface oil vaporizes or in diesel engines even burns off. This vaporized oil constitutes the majority of the oil consumption by modern-day internal-combustion engines.

The special shape of the piston rings is designed to facilitate the build-up of an oil wedge. A critical factor is, on the one hand, the piston speed, which must not be allowed to become too high, as otherwise the shear forces can no longer be absorbed by the oil and the oil film will break. On the other hand, the reversal points, particularly at top dead center, must be classed as critical with regard to wear.

Empirical values and data for calculation

Characteristics

Fuel consumption

Figure 1 shows in a diagram nine different specific consumptions plotted against the effective mean pressure p_{me}. Since p_{me} has a proportional effect on the equation of energy ($E_E = i p_{me} V_H$; $i = 1$ for two-stroke engines, $i = 0.5$ for four-stroke engines), the absolute consumption rises with increasing p_{me}. This situation is illustrated for example by straight line 1.

The specific consumptions rise from straight lines 1 through 9. In other words, for the first process 200 g of fuel are needed to obtain 1 kWh of energy. For 9, 500 g of fuel already are expended for 1 kWh.

The y-axis is normalized to the working cycle and the displacement (swept volume). In other words, the energy E_E is directly dependent on p_{me}. When p_{me} increases, so too the energy increases. More fuel is needed when the energy increases. If a certain p_{me} is taken – e.g. 15 bar, which corresponds to a defined energy – significantly more fuel is required for straight line 5 than for straight line 1 (roughly 120 mg as against 80 mg) to generate the same amount of mechanical energy. This is also discernible from the gradient of the straight line.

Characteristic values of gasoline and diesel engines

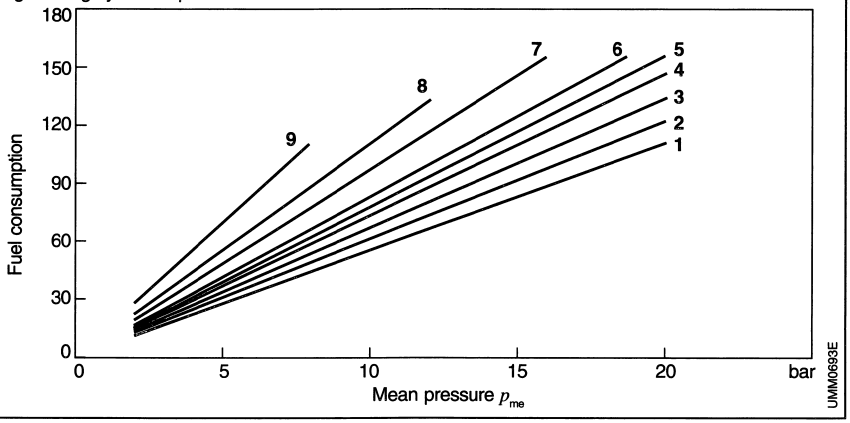
Table 1 shows characteristic values in a comparison of gasoline and diesel engines. What is noticeable at the outset is the large differences in rpm between gasoline and diesel engines. The compression ratios too are – by the nature of the process – much higher for diesel engines than for gasoline engines. Due to the higher compression pressure, greater mean pressures p_{me} are possible. Since diesel engines attain lower speeds (rpm), they have a lower power output per liter (power output per displacement). Due to the lower power output per liter and the resulting higher weight for durability reasons, diesel engines have the higher weight-to-power ratio (vehicle mass per power output). Due to the lower compression ratio, gasoline engines have an unfavorable specific consumption.

Torque position

The torque position denotes the position on the engine-speed curve (Figure 1) at which maximum torque ($n_{Md,max}$) is developed, relative to the rated-power speed

Figure 1: Fuel consumption

1 200 g/kWh, 2 220 g/kWh, 3 240 g/kWh,
4 260 g/kWh 5 280 g/kWh, 6 300 g/kWh,
7 350 g/kWh, 8 400 g/kWh, 9 500 g/kWh.



(n_{nom}) in percent ($n_{Md,max}/n_{nom}$). For typical values, see Table 2.

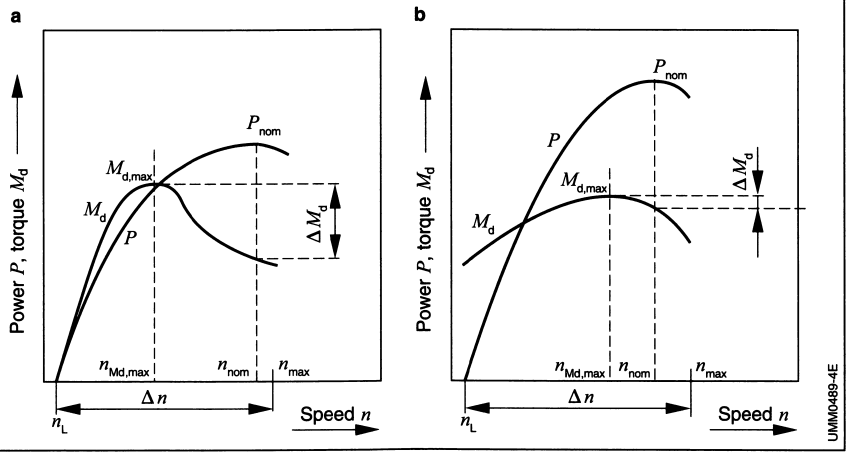
Useful speed range
The useful speed range Δn shows the range in which an engine is operated. For example, a gasoline engine with a maximum speed n_{max} of 7000 rpm and an idle speed of 800 rpm has a useful speed range of 6200 rpm. The figures for useful speed range in Table 2 are given in rpm x 1000.

Table 1: Comparative data

	Engine type and application		Engine speed n_{nom} rpm	Compression ratio ε	Max. mean pressure p_{me} bar	Power output per liter kW/l	Weight-to-power ratio kg/kW	Specific fuel consumption g/kWh	Torque increase %
Gasoline engines	Motorcycles	4-stroke	5000...13000	9...12	9...13	50...150	2.5...0.5	230...280	10...15
	Pass. cars	NAE	5000...8000	9...13	11...14	40...80	2.0...0.8	220...270	15...20
		SCE/IC	5000...7500	9...12	15...22	60...110	1.5...0.5	220...250	20...40
Diesel engines	Pass. cars, light comm. veh.	SCE/IC	3500...4500	18...22	12...20	35...55	3.0...1.3	200...220	20...40
	Comm. veh.	SCE/IC	1800...2600	15...18	18...24	25...40	4.0...2.5	180...210	20...40

NAE Naturally aspirated engine; SCE Supercharged engine; IC Intercooling (charge-air cooling)

Figure 2: Typical power and torque curves
a) Diesel engine with exhaust-gas turbocharging,
b) Gasoline engine (naturally aspirated engine).
n_L Idle speed, $n_{Md,max}$ Engine speed at $M_{d,max}$, n_{nom} Nominal speed, n_{max} Maximum engine speed, Δn Useful speed range,
P_{nom} Rated power, $M_{d,max}$ Maximum torque.



Torque rise
The torque rise ΔM_d denotes the difference between the maximum torque $M_{d,max}$ and the torque at the rated-power point. The values in Table 3 are given in percent.

Torque increase
The torque increase is the gradient of the torque curve between maximum torque and the torque at rated power.

Engine output and atmospheric conditions
The torque and thus the power output of an internal-combustion engine are essentially determined by the heat content of the cylinder charge. The flow rate of air (or, more precisely, of oxygen) in the cylinder charge provides a direct indication of heat content. The change which the engine will display at full power can be calculated as a function of variations in the condition of the ambient air (temperature, barometric pressure, humidity), provided that engine speed, air/fuel (A/F) ratio, volumetric efficiency, combustion efficiency, and total engine power loss remain constant. The A/F mixture responds to lower atmospheric density by becoming richer. The volumetric efficiency (ratio of pressure in cylinder at BDC to pressure in ambient air) only remains constant for all atmospheric conditions at maximum throttle-valve aperture (WOT). Combustion efficiency drops in cold thin air as vaporization rate, turbulence, and combustion speed all fall. Engine power loss (friction losses + gas-exchange work + boost power drain) reduces the indicated power.

Table 2: Typical values for useful speed range and torque position

Engine type		Useful speed range Δn_N (×1000 rpm)	Torque position [%]
Diesel engine	Pass. cars	3.5...5	15...40
	Comm. veh.	1.8...3.2	10...60
Gasoline engine		4...7	25...35

Table 3: Typical values for torque rise

Engine type		Torque rise ΔM_d [%]
Diesel engine, pass. cars	SCE and IC	25...35
Diesel engine Comm. veh.	SCE and IC	25...40
Gasoline engine	Nat. asp. engine	25...30
	SCE and IC	30...35
SCE Supercharged engine; IC Intercooling (charge-air cooling)		

Effect of atmospheric conditions
The quantity of air which an engine draws in, or is inputted to the engine by supercharging, depends upon ambient-air density; colder, heavier, denser air increases engine output. As a rule of thumb: Engine power drops by approximately 1 % for each 100 m rise in elevation. Depending upon engine design, the cold intake air is normally heated to some degree while traversing the intake passages, thereby reducing its density and thus the engine's ultimate output. Humid air contains less oxygen than dry air and therefore produces lower engine power outputs. The decrease is generally modest to the point of insignificance. The warm humidity of air in tropical regions can however result in a noticeable engine power loss.

Definitions of power
The effective power is the engine's power as measured at the crankshaft or ancillary mechanism (such as the transmission) at the specified engine speed. When measurements are made downstream from the transmission, the transmission losses must be factored into the equation.

The rated power is the maximum effective power of the engine at full throttle. The net power corresponds to the effective power.

Conversion formulas are used to convert the results of dynamometer testing to reflect standard conditions, thereby negating the influences of such factors as time of day and year while simultaneously allowing the various manufacturers to provide mutually comparable data. The procedure converts air density – and thus the effective volume of air in the engine – to defined "standard conditions" for air mass.

Table 4: Power-output correction factors

Gasoline engines, naturally aspirated and supercharged			
	EEC 80/1269 [9], JIS D 1001 [10], SAE J 1349 [11]		DIN 70020 [12]
Correction factor α_a	$\alpha_a = A^{1.2} \cdot B^{0.5}$ $A = 99/p_{PT}$ $B = T_p/298$		$\alpha_a = A \cdot B^{0.5}$ $A = 101.3/p_{PF}$ $B = T/293$
Corrected power: $P_0 = \alpha_a P$ (kW) (P measured power)			
Diesel engines			
Atmospheric engine correction factor f_a	$f_a = A^{0.7} \cdot B^{1.5}$ ($A = 99/p_{PT}$; $B = T_p/298$) (turbocharged engines with/without intercooling).		As α_a for gasoline engines
Engine correction factor f_m	$40 \le q/r \le 65$: $q/r < 40$: $q/r > 65$:	$f_m = 0.036\,(q/r) - 1.14$ $f_m = 0.3$ $f_m = 1.2$	$f_m = 1$
$r = p_L/p_E$ Charge-air pressure response, with p_L absolute charge-air pressure, p_E Absolute pressure upstr. of compressor, q Spec. fuel consumption (SAE J 1349, [8]), p_{PT} Barometer pressure dry, p_{PF} Barometer pressure absolute, T_P Test-dynamometer temperature. 4-stroke engines: $q = 120{,}000\ F/DN$; 2-stroke engines: $q = 60{,}000\ F/DN$, with F Fuel flow [mg/s], D Displacement [l]; N Engine speed [rpm].			
Corrected power: $P_0 = P f_a^{fm}$ (kW) (P measured power).			

The comparison data in Table 4 show the most important factors for these power-output corrections.

Displacement

The displacement – also referred to as the swept volume – is characterized by the total vertical stroke of the piston multiplied by the cylinder diameter. The displacement of an individual cylinder is characterized as V_h. The displacement V_H of an engine corresponds to the sum total of all the displacements V_h of its cylinders.

Compression volume

The compression volume V_c corresponds to the compression space in the cylinder when the piston is at TDC.

Combustion chamber

The combustion chamber is obtained from the sum total of displacement and compression volume.

Compression ratio

The compression ratio describes the ratio of the combustion-chamber volume when the piston is at BDC to the volume when the piston is at TDC.

Piston stroke

The piston stroke is a function of the crank angle. It is not uniform over one revolution and varies. Thus, the piston moves from TDC towards BCD up to 90 °CA after TDC much faster downwards than from 90 ° to BDC. The mean piston speed is the average speed of the piston stroke over one complete revolution.

Efficiency

The efficiency describes the ratio of the mechanically supplied power to the energy supplied chemically by the fuel.

Calculations

Table 6 shows calculation formulas for some quantities in the engine. The quantity equations are universally valid, the numerical values in the given unit are inserted in the numerical-value equations. These equations tend to be applied in practice.

The frequently used quantities and their units are set out in Table 5.

Table 5: Quantities and units

Quantity		Unit
a_K	Piston acceleration	m/s²
B	Fuel consumption	kg/h; dm³/h
b_e	Spec. fuel consumption	g/kWh
d	Cylinder diameter	mm
d_v	Valve diameter	mm
F	Force	N
F_G	Gas force in the cylinder	N
F_N	Piston side thrust	N
F_o	Oscillating inertial force	N
F_r	Rotating inertial force	N
F_s	Connecting-rod force	N
F_T	Tangential force	N
M	Torque	Nm
M_o	Oscillating moments	Nm
M_r	Rotating moments	Nm
M_d	Engine torque	Nm
m_p	Weight-to-power ratio	kg/kW
n	Engine speed	rpm
n_p	Fuel-injection pump speed	rpm
P	Power	kW
P_{eff}	Effective power	kW
P_H	Power output per liter	kW/dm³
p	Pressure	bar
p_c	Final compression pressure	bar
p_e	Mean piston pressure (mean pressure, mean working pressure)	bar
p_L	Charge-air pressure	bar
p_{max}	Peak injection pressure in the cylinder	bar
r	Crankshaft radius	mm
s_d	Injection cross-section of the nozzle	mm²
S, s	Stroke, general	mm
s	Piston stroke	mm
s_f	Induction stroke of a cylinder (2-stroke)	mm
s_F	Induction stroke, 2-stroke engine	mm
S_k	Piston clearance from TDC	mm
S_s	Slot height, 2-stroke engine	mm
T	Temperature	°C, K
T_c	Final compression temperature	K
T_L	Charge-air temperature	K
T_{max}	Peak temperature in combustion chamber	K
t	Time	s
V	Volume	m³

Quantity		Unit
V_c	Compression volume of a cylinder	dm³
V_E	Injected fuel quantity per pump stroke	mm³
V_f	Charge volume of a cylinder (2-stroke)	dm³
V_F	Charge volume of a 2-stroke engine	dm³
V_h	Displacement of a cylinder	dm³
V_H	Displacement of the engine	dm³
v	Velocity	m/s
v_d	Mean velocity of injection spray	m/s
v_g	Gas velocity	m/s
v_m	Mean piston speed	m/s
v_{max}	Max. piston speed	m/s
z	Number of cylinders	–
β	Pivoting angle of connecting rod	°
ε	Compression ratio	–
η	Efficiency	–
η_e	Net efficiency	–
η_{th}	Thermal efficiency	
v, n	Polytropic exponent of real gases	–
ρ	Density	kg/m³
φ, α	Crank angle (φ_o = top dead center)	°
ω	Angular velocity	rad/s
λ	= r/l Stroke/connecting-rod ratio	–
λ	Excess-air factor	–
κ	= c_p/c_v Adiabatic exponent of ideal gases	–

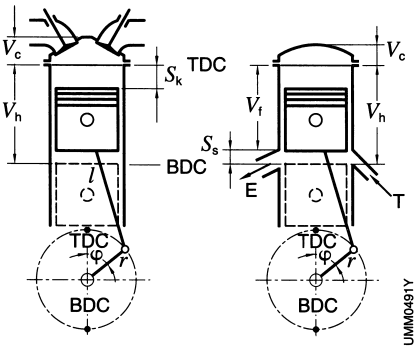
Superscripts and subscripts

0, 1, 2, 3, 4, 5	Cycle values/main values
o	Oscillating
r	Rotating
1st, 2nd	1st, 2nd order
A	Constant

Conversion of units

1 g/hp·h	= 1.36 g/kWh
1 g/kWh	= 0.735 g/hp·h
1 kpm	= 9.81 Nm ≈ 10 Nm
1 Nm	= 0.102 kpm ≈ 0.1 kpm
1 hp	= 0.735 kW
1 kW	= 1.36 hp
1 at	= 0.981 bar ≈ 1 bar
1 bar	= 1.02 at ≈ 1 at

Table 6: Calculation equations

Mathematical relationship between quantities	*Numerical relationship between quantities*
Displacement Displacement of a cylinder $V_h = \frac{\pi d^2 s}{4}$ (4-stroke); $V_f = \frac{\pi d^2 s_f}{4}$ (2-stroke) Displacement of the engine $V_H = V_h z$ (4-stroke); $V_F = V_f z$ (2-stroke)	$V_h = 0.785 \cdot 10^{-6} d^2 s$ V_h in dm³, d in mm, s in mm $V_H = 0.785 \cdot 10^{-6} d^2 s z$ V_h in dm³, d in mm, s in mm
Compression Compression ratio $\varepsilon = \frac{V_h + V_c}{V_c}$ Final compression pressure $p_c = p_o \varepsilon^v$ Final compression temperature $T_c = T_o \varepsilon^{v-1}$	 4-stroke engine 2-stroke engine $\varphi = 6\, n\, t$ φ in °, n in rpm, t in s
Piston stroke Piston clearance from top dead center $S_k = r\left[1 + \frac{l}{r} - \cos\varphi - \sqrt{\left(\frac{l}{r}\right)^2 - \sin^2\varphi}\right]$ Crank angle $\varphi = 2\,\pi\, n\, t$ (φ in rad) Piston speed (approximation) $v \approx 2\pi n r\left(\sin\varphi + \frac{r}{2l}\sin 2\varphi\right)$ Mean piston speed $v_m = 2ns$ Maximum piston speed (approximate, if connecting rod is on a tangent with the big-end trajectory; $a_k = 0$	$v \approx \frac{ns}{19100}\left(\sin\varphi + \frac{r}{2l}\sin 2\varphi\right)$ v in m/s, n in rpm, l, r and s in mm $v_m = \frac{ns}{30000}$ v_m in m/s, n in rpm, s in mm
Piston acceleration (approximation) $a_k \approx 2\pi^2 n^2 s\left(\cos\varphi + \frac{r}{l}\cos 2\varphi\right)$	$a_k \approx \frac{n^2 s}{182400}\left(\cos\varphi + \frac{r}{l}\cos 2\varphi\right)$ a_k in m/s², n in rpm, l, r and s in mm

l/r	3.5	4	4.5
v_{max}	1.63 v_m	1.62 v_m	1.61 v_m

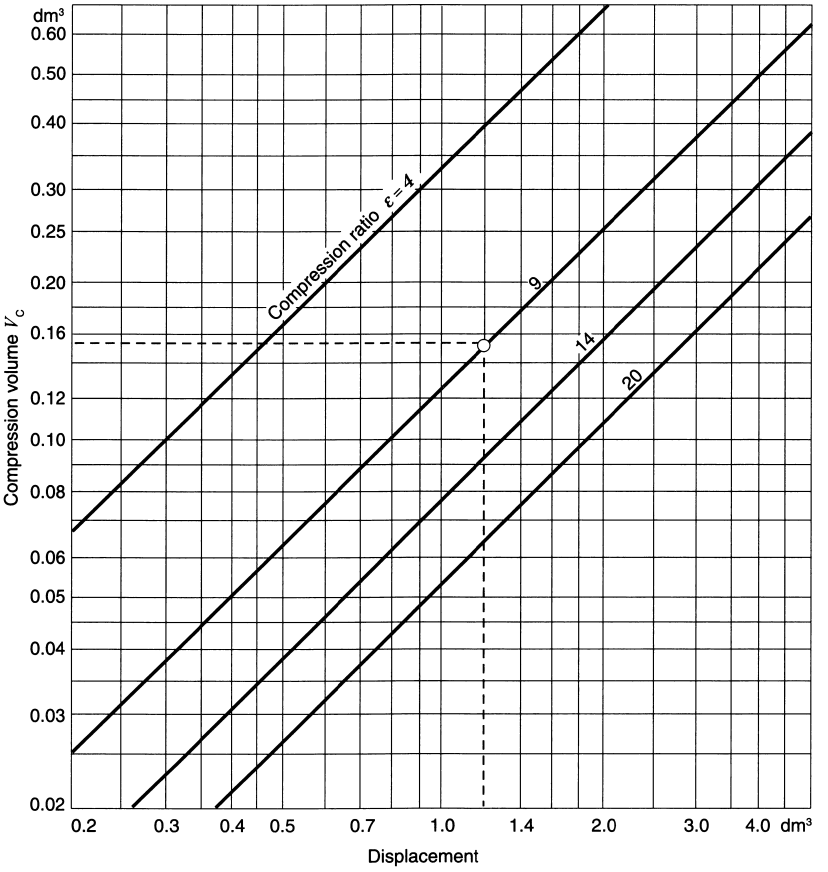
Table 6: Calculation equations (continued)

Mathematical relationship between quantities	*Numerical relationship between quantities*
Gas velocity Mean gas velocity in the valve section $v_g = \frac{d^2}{d_v^2} v_m$	$v_g = \frac{d^2}{d_v^2} \cdot \frac{ns}{30000}$ v_g in m/s, d, d_v, and s in mm, n in rpm
Engine power $P = M\omega = 2\pi Mn$ $P_{eff} = V_H p_e \frac{n}{K}$ $K = 1$ for 2-stroke engine $K = 2$ for 4-stroke engine Power per unit displacement (power output per liter) $P_H = \frac{P_{eff}}{V_H}$ Weight-to-power ratio $m_p = \frac{m}{P_{eff}}$	$P = \frac{Mn}{9549}$ P in kW, M in N m (= W s), $P_{eff} = \frac{V_H p_e n}{K \cdot 600} = \frac{M_d n}{9549}$ P_{eff} in kW, p_e in bar, n in rpm M_d in N m $P = \frac{Mn}{716.2}$ P in hp, M in kp · m, n in rpm

Table 6: Calculation equations (continued)

Mathematical relationship between quantities		*Numerical relationship between quantities*	
Mean piston pressure (mean pressure, mean working pressure)			
4-stroke engine $p = \frac{2P}{V_H n}$	2-stroke engine $p = \frac{2P}{V_H n}$	4-stroke engine $p = 1200 \frac{P}{V_H n}$	2-stroke engine $p = 600 \frac{P}{V_H n}$
		p in bar, P in kW, V_H in dm^3, n in rpm	
		$p = 833 \frac{P}{V_H n}$	$p = 441 \frac{P}{V_H n}$
		p in bar, P in hp, V_H in dm^3, n in rpm	
$p = \frac{4\pi M}{V_H}$	$p = \frac{2\pi M}{V_H}$	$p = 0.1257 \frac{M}{V_H}$	$p = 0.0628 \frac{M}{V_H}$
		p in bar, M in N m, V_H in dm^3	
Engine torque			
$M_d = \frac{V_H P_e}{4\pi}$	$M_d = \frac{V_H P_e}{4\pi}$	$M_d = \frac{V_H P_e}{0.12566}$	$M_d = \frac{V_H P_e}{0.06284}$
		M_d in N m, V_H in dm^3, p_e in bar	
		$M_d = 9549 \frac{P_{eff}}{n}$	
		M_d in N m, P_{eff} in kW, n in rpm	
Fuel consumption			
B = Measured values in kg/h $b_e = \frac{B}{P_{eff}}$ $b_e = \frac{1}{H_u \eta_e}$		B in dm^3/h or kg/h V_B = Measured volume on test dynamometer t_B = Elapsed time for measured volume consumption $b_e = 3600 \frac{V_b \rho_B}{t_b P_{eff}}$ ρ_B = Fuel density in g/cm^3, t_B in s, V_B in cm^3, P_{eff} in kW.	
Efficiency			
$\eta_{th} = 1 - \varepsilon^{1-k}$ $\eta_e = \frac{P_{eff}}{B H_u)}$		$\eta_e = \frac{x}{b_e}$ $x = 82$ for $H_u = 44$ $x = 86$ for $H_u = 42$ $x = 90$ for $H_u = 40$ $x = 120$ for $H_u = 30$ H_u Specific calorific value in MJ/kg b_e Specific fuel consumption in g/kWh	

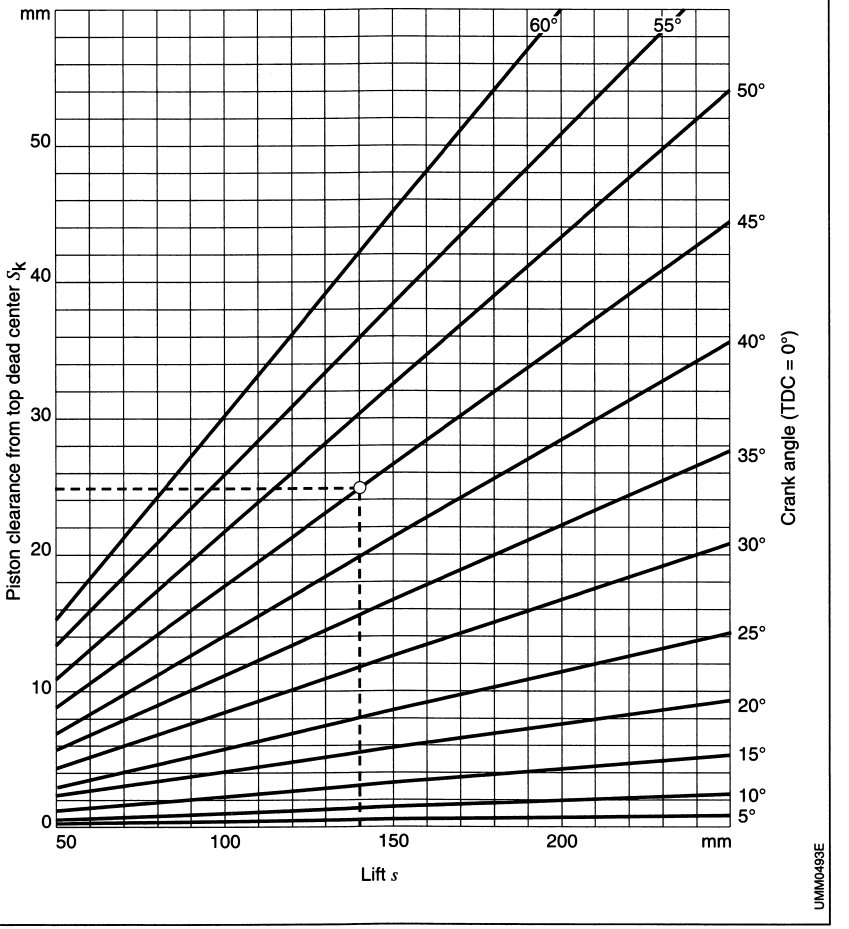
Figure 3: Displacement and compression space
The diagram applies to the displacement V_h and the compression space V_c of the individual cylinder and to the total displacement V_H and the total compression space V_C.



Displacement and compression space
In Figure 3 the displacement is plotted on the x-axis and the compression space is plotted on the y-axis. The straight lines show lines with a constant compression ratio. As the displacement increases, so too the compression space increases. If the compression ratio is raised, the compression space must become smaller for the same displacement.

An engine with a displacement of 1.2 dm³ and a compression ratio ε = 9 has a compression space of approximately 0.15 dm³.

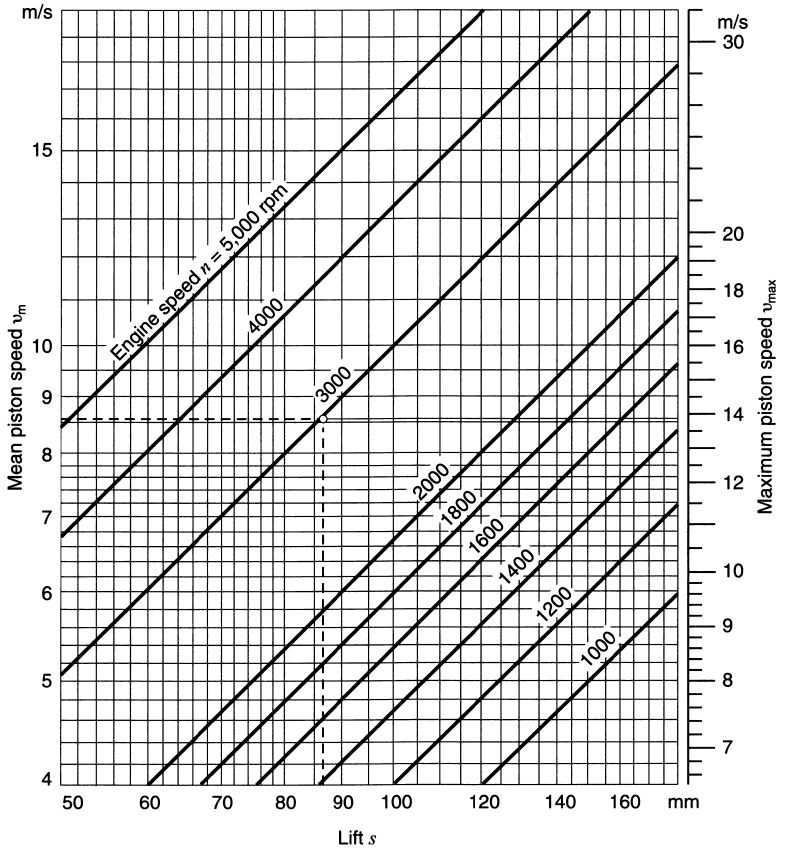
Figure 4: Piston clearance from top dead center
Conversion of degrees of crankshaft angle to mm piston travel.



Determining the piston clearance
Example
For a stroke of 140 mm the piston clearance from top dead center at 45 °CA is 25 mm (Figure 4).

The diagram is based on a crank ratio $l/r = 4$ (l connecting-rod length, r on half of the stroke length). However, it applies with good approximation (error less than 2 %) to all crank ratios l/r between 3.5 and 4.5.

Figure 5: Piston speed



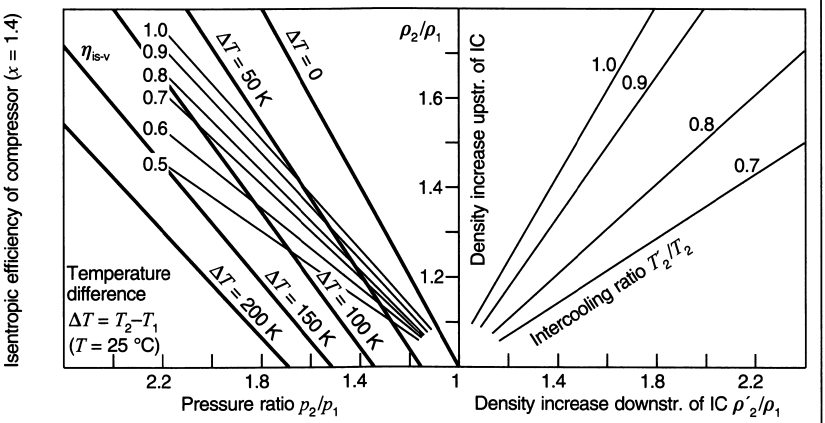
Determining the piston speed
The x-axis in Figure 5 shows the stroke in millimeters. The left y-axis shows the mean and the right y-axis the maximum piston speed. In addition, lines of constant speed are marked. It can clearly be seen that both piston speeds increase as the stroke increases. The same applies to increasing speeds.

Example
For the stroke s = 86 mm and at an engine speed of n = 3000 rpm the mean piston speed v_m = 8.7 m/s and the maximum piston speed
v_{max} = 13.8 m/s.

The diagram is based on $v_{max} = 1.62\, v_m$.

Figure 6: Density increase of combustion air in cylinder on turbo/supercharging

Increase in density on supercharging as a function of the pressure ratio in the compressor, of the compressor efficiency, and of the intercooling rate for intercooling (IC).

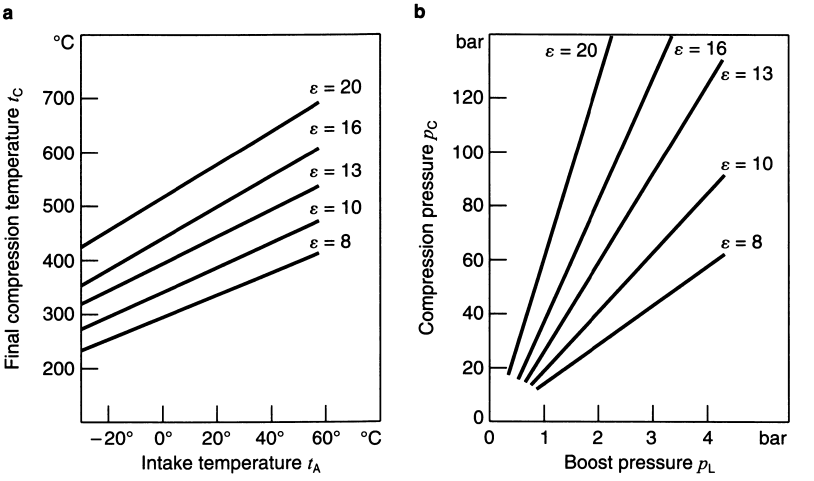


$p_2/p_1 = \pi_c$ = Pressure ratio during precompression
ρ_2/ρ_1 = Increase in density, ρ_1 = Density upstream of compressor, ρ_2 = Density downstream of compressor in kg/m³
T_2'/T_2 = Intercooling rate, T_2 = Temperature before IC, T_2' = Temperature after IC in K
η_{is-v} = Isentropic compressor efficiency

UMM0495-3E

Figure 7: Final compression pressure and final compression temperature

a) Final compression temperature as a function of compression ratio and intake temperature,
b) Final compression pressure as a function of compression ratio and charge-air pressure.



$t_c = T_c - 273{,}15$ K, $T_c = T_A\, \varepsilon^{n-1}$, $n = 1{,}35$

$p_c = p_L\, \varepsilon^n$, $n = 1.35$

UMM0496E

Determining the density increase of the combustion air

Figure 6 shows that as the pressure ratio increases, depending on how strongly the air is heated and cooled in the intercooler, the density of the gas increases. The hotter the gas after the intercooler, the smaller the density increase.

Determining the final compression pressure

Figure 7a shows on the x-axis the intake temperature and on the y-axis the final compression temperature. In addition, lines of a constant compression ratio are marked. It can be seen that the final compression temperature increases as the intake temperature increases. The same applies to increasing compression ratios, where here the straight-line gradient increases with a higher compression ratio. This graphic shows why lowering the compression ratio can reduce the knock tendency.

Figure 7b shows the final compression pressure as a function of compression ratio and charge-air pressure. If both increase, the final compression pressure also increases. However, a significant effect on the final compression pressure can be seen by increasing the compression ratio.

References

[1] H. Hütten: Motoren – Technik, Praxis, Geschichte. Motorbuch-Verlag, Stuttgart, 1997.
[2] DIN 73021: Designation of the rotational direction, the cylinders and the ignition circuits of motorcar engines.
[3] ISO 1204: Reciprocating internal combustion engines – Designation of the direction of rotation and of cylinders and valves in cylinder heads, and definition of right-hand and left-hand in-line engines and locations on an engine.
[4] I. Bronstein; K. Semendjajew: Taschenbuch der Mathematik. 7th Edition, Verlag Harri Deutsch, 2008.
[5] DIN 4760: Form deviations; Concepts, Classification system.
[6] DIN EN ISO 13565: Geometrical Product Specifications (GPS) – Surface texture: Profile method – Surfaces having stratified functional properties.
[7] DIN EN ISO 11562: Geometrical Product Specifications (GPS) – Surface texture: Profile method – Metrological characteristics of phase correct filters.
[8] SAE J 1349: Engine Power Test Code Spark Ignition and Compression ignition As Installed Net Power Rating.
[9] Council Directive 80/1269/EEC of 16 December 1980 on the approximation of the laws of the Member States relating to the engine power of motor vehicles.
[10] JIS D 1001: Road vehicles - Engine power test code.
[11] SAE J 1349: Engine Power Test Code Spark Ignition and Compression ignition As Installed Net Power Rating.
[12] DIN 70020: Road vehicles – Automotive engineering.

Exhaust-gas emissions

Historic measures

Endeavors at controlling air pollution are an absolutely integral element of environmental protection. It successes since the mid-20th century are thanks in particular to the economic analysis of the generation of pollutants and the derived further development of the combustion process, the consistent improvement of fuel quality, and the technical development of exhaust-gas treatment.

Smog and acid rain are still today common catchwords that have their origin in the second half of the 20th century. On region hugely affected by smog in that period was the US state of California. In order to reduce air pollution in areas of high population density such as Los Angeles, California had already enacted the first laws in the early 1960s (see [1]). Other states followed California's example. In 1970 the European Community for its part passed its first emission regulations. These legislative measures induced automobile manufacturers to fit their motor vehicles with exhaust-gas treatment systems in the markets affected. This has resulted in today's global market for exhaust-gas treatment systems.

Constituents of engine exhaust gas

Major constituents

Various combustion products are produced in internal-combustion engines during the combustion of fuel. Exhaust gas comprises primarily the non-toxic major constituents:

- nitrogen N_2 (constituent of the intake air),
- water vapor H_2O,
- carbon dioxide CO_2,
- oxygen O_2 (in diesel engines and gasoline engines with lean-burn operation).

Carbon dioxide is present in the air as a natural constituent and is not categorized as a pollutant in respect of exhaust-gas emissions for motor vehicles. However, it is considered to be one of the causes of the greenhouse effect and the associated global climate change. Since industrialization, the CO_2 content in the atmosphere has risen by roughly 40 % to today's figure of 400 ppm.

The amount of released carbon dioxide is directly proportional to fuel consumption. Measures taken to reduce fuel consumption are therefore becoming increasingly significant.

Pollutants

The above-mentioned constituents are the components of exhaust gas that would be produced in a complete, ideal combustion of pure fuel – i.e. in the complete combustion of fuel with oxygen without unwanted secondary reactions. However, minor constituents are also produced in real vehicle operation.

The quantity of minor constituents produced during combustion is very much dependent on the engine operating status. For the gasoline engine, the quantity of untreated exhaust gas (exhaust gas after combustion and before exhaust-gas treatment) amounts to approximately 1 % of the total amount of the exhaust gas for an engine at normal operating temperature and with a stoichiometric mixture composition (where $\lambda = 1$). The compo-

sition of diesel exhaust gas is very much dependent on the excess air.

The essential pollutants produced directly during the in-engine combustion of fuels are described in the following.

Hydrocarbons
Fuels that contain hydrocarbons never combust absolutely completely in engines. The incompletely combusted hydrocarbons emitted from an engine and their concentrations are determined in particular by the combustion and engine applications with injection strategy, injected fuel quantity, injection pressure, point of injection, etc., and by the air-fuel ratio λ, by the combustion temperature, and by further engine influencing factors.

Aliphatic hydrocarbons (alkanes, alkenes, alkines as well as their cyclic derivatives) are very nearly odorless. Cyclic aromatic hydrocarbons (e.g. benzene, toluene, polycyclic hydrocarbons) exude an odor. Partially oxidized hydrocarbons (e.g. aldehydes, ketones) have an unpleasant odor and form derived products when exposed to sunlight.

Carbon monoxide
Carbon monoxide is formed during the incomplete combustion of fuel due to an insufficient oxygen supply. This can be caused by locally rich areas as a result of insufficient mixture preparation (e.g. unvaporized fuel droplets or fuel film on the wall of the combustion chamber).

Carbon monoxide is a colorless, odorless, and tasteless gas. It is toxic since it binds to the hemoglobin in blood and thereby inhibits oxygen transport.

Particulates
Unlike gaseous exhaust-gas constituents, particulates are not a clearly defined chemical species. All solid or liquid combustion residues which can accumulate in a defined sampling process on a sample filter are called particulates. These are, aside from soot particles – consisting of a solid soot core and adherent incompletely combusted hydrocarbons resulting from fuel and lube oil – droplets of hydrocarbons, water and sulfurous acids (sulfates) as well as ash particles and abraded metal.

Sulfur oxides
The term sulfur oxides (SO_x) denotes a mixture of sulfur dioxide (SO_2) and sulfur trioxide (SO_3). Sulfur oxides are almost inevitably formed during the combustion of fuel containing sulfur. The sulfur content in the fuel must always be taken into consideration in the choice and design of exhaust-gas treatment systems in order to guarantee their long-term effectiveness.

Nitrogen oxides (nitric oxides)
The nitrogen content in fuels is generally low and plays a subordinate role in the formation of nitrogen oxides. During the combustion of fossil fuels nitrogen oxides are produced from reactions of intermediately formed hydrocarbon radicals and oxygen (O_2) with atmospheric nitrogen (N_2). Phenomenologically these reactions have hitherto not been explained in every detail despite broad empirical knowledge. The layout of the combustion chamber and the combustion application have a large influence, e.g. guidance of the flame front and its local temperatures.

The produced mixture of strongly oxidizing nitrogen oxides is termed "nitrous gases" and comprises nitrogen monoxide (NO), nitrogen dioxide (NO_2) and various dimers (e.g. N_2O_3, N_2O_4) which are in thermodynamic equilibrium. Laughing gas (N_2O) is not strongly oxidizing and is not classified as a nitrous gas. It is a strong greenhouse gas. When it comes to the frequently used generic term NO_x, it is always necessary to define whether laughing gas is included or not. Both definitions are common.

Basic principles of catalytic exhaust-gas treatment

It seems advantageous to convert the pollutants formed in the engine with each other or with another substance contained in the exhaust gas into less harmful substances. The following challenges must be surmounted for this purpose:
- In the fewest engine operating states the exhaust gas contains equivalent proportions of the substances to be able to convert these with each other.
- The exhaust-gas temperature is too low at many engine operating points to be able to get over the activation barrier of the desired reactions.

The strategy of catalytic exhaust-gas treatment starts at this point.

Catalytic exhaust-gas treatment
A catalyst is scientifically defined as a substance in the presence of which a certain reaction or a reaction sequence takes place more quickly under kinetically limited reaction conditions and which itself emerges unchanged. In other words, the catalyst reduces the activation barrier of the reaction catalyzed by it. However, the components of exhaust-gas treatment systems designated in the technical sense as a catalyst furthermore perform further functions (see Exhaust-gas treatment, Catalytic converter).

Catalyst components
Catalysts for exhaust-gas treatment consist of complex-structured, highly porous ceramics which are in the main applied as coatings (catalyst coating) to honeycombed carrier structures. A distinction is made between the following catalyst components.

Catalytically active components
Catalytically active components (e.g. noble metals) increase the speed of the desired reactions under suitable conditions.

Storage components
Storage components (e.g. zeolites) absorb certain substances from the exhaust gas and store them until suitable conditions exist to convert them.

Carrier components
Carrier components (e.g. Al_2O_3) form a ceramic matrix in the large inner surfaces of which catalytically active components are present in finely dispersed form.

Stabilizing components
Stabilizing components protect the carrier and storage components as well as the catalytically active components against excessive thermal and chemical impairments.

Honeycombs
The honeycombs to be coated in the main exhibit 10 to 250 parallel channels per square centimeter. They are usually extruded from a ceramic, e.g. cordierite, or manufactured as a special form from metal foil.

Arrangement in the exhaust-system branch
The catalysts are preferably integrated into the exhaust-gas system in such a way that the gases flow through them as uniformly as possible. Each deviation from the ideal distribution of the gases reduces the theoretically possible conversion of a catalyst such that the nonuniform flow necessitates an accordingly more powerful, i.e. larger catalyst for example.

The volume of the honeycombs, which in the technical sense is termed the catalyst volume, depends in particular on the required quantity of catalytic components which is required for the desired reduction of pollutants. A limiting factor in this respect is often the quantity which can be applied per unit of volume to the honeycomb. On the engine side the exhaust-gas back pressure of the coated honeycomb constitutes a limitation. On the vehicle side the available installation space frequently limits the catalyst volume. In Euro 6 exhaust-emissions standard diesel vehicles the total catalyst volume is often three to six times the engine displacement.

Catalytic processes for exhaust-gas treatment

Boundary conditions

Catalysts (catalytic converters) are the basis of modern processes for treating the exhaust gases of motor vehicles. Boundary conditions must be satisfied to guarantee their effectiveness in the long term. For example, the quality of the operating fluids must be matched to the respective catalysts. Leaded fuel can cause a rapid deactivation of the catalysts. Sour fuel (i.e. containing sulfur), heavy metals from the engine oil, inorganic additives, and other not completely combusted substances in the engine can cause progressive contamination of the catalysts quickly or slowly over the vehicle's service life. In this way, the efficiency of the exhaust-gas treatment system is drastically reduced in part.

In the case of the catalyst requirements imposed on the engine and vehicle applications it is also essential to adhere to certain exhaust-gas composition limits and the exhaust-gas temperature, even if critical conditions in some cases only occur on a limited basis in terms of time and space on the catalyst surface. For example, temperatures which lead to local fusing and thus destruction of a particulate filter are to be avoided during soot burn-off. This necessitates, for example, a limitation of the amount of soot to be stored.

Generally speaking, a faster progression of the catalyzed reaction can be achieved in a catalytic process with a larger quantity of active components – if the activation barrier is exceeded and if substance-transport processes do not limit the reaction speed. But even when large quantities are used the exhaust-gas temperature is always too low directly after engine starting to achieve significant pollutant conversion. The catalyst heats up with a constant exhaust-gas flow. The temperature at which 50 % of the pollutant emitted by the engine is converted is called the light-off temperature. During cold starting of a diesel engine several minutes regularly elapse before the catalysts achieve optimum operating conditions. This is counteracted by applications with engine heating measures. Electrical heating measures have also been implemented in special applications. Storage components of the catalysts converter moreover store pollutants in the cold-start phase and desorb them once the light-off temperature is exceeded. An example: Special zeolites are suitable for storing some hydrocarbons in such a way.

References

[1] https://www.arb.ca.gov/html/brochure/history.htm

Belt drives

Friction belt drives

Application

Friction belt drives are used in automobiles predominantly to drive accessories (Figure 1). These drives had previously been driven primarily by V-belts. However, the significantly higher power density required in today's applications, e.g. due to space restrictions and higher power consumption of the accessories, have resulted in these having been replaced in the meantime almost exclusively by serpentine drives with ribbed V-belts (Micro-V® belts). They are typically used to drive the alternator, the A/C compressor, the power-steering pump, but also fans, mechanical chargers, or pumps for secondary air injection.

Ribbed V-belts are also used to start engines in hybrid or in start/stop systems. Here, the conventional starter is replaced by a starter alternator, and the belt transmits the starter torque to the crankshaft to start the engine. As well as starting, supplementary functions such as additional engine acceleration (boost mode) and energy recovery by engine braking (recuperation) can be realized.

Forces and loads in the belt drive

The transmittable power is determined from (quantities, see Table 1)

$$P = (F_1 - F_2) \cdot \frac{v}{1000}, \; P \text{ in kW.}$$

Table 1: Symbols and units

Quantity	Unit
F_1 Belt-run force on load side	N
F_2 Belt-run force on slack side	N
F_U Peripheral force	N
F_{HL} Pretensioning force or "hubload"	N
F_B Loading of belt by wrap	N
F_C Loading of belt by centrifugal forces in wrap angle	N
M Transmitted torque	Nm
v Belt speed	m/s
P Required power transmission	kW
α_R Groove angle, belt	°
α_S Groove angle, pulley	°
β Wrap angle	rad
μ Coefficient of friction between belt rib and pulley groove	–
μ' Coefficient of friction referred to smooth pulley	–
ω Angular frequency	s^{-1}
L_B Reference belt length	mm
U_B Reference circumference	mm

Figure 1: Application example, accessory drive

1 Alternator, 2 Tension pulley, 3 Idler pulley, 4 Crankshaft, 5 Water pump, 6 Power-steering pump, 7 A/C compressor.

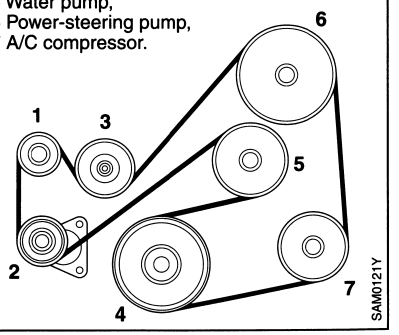
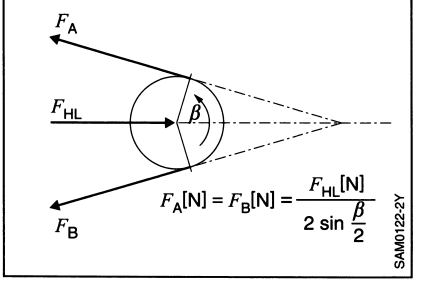


Figure 2: Forces on an unloaded belt pulley

β Wrap angle, F_A Belt-run force, F_{HL} Hubload, F_B Belt-run force.



The transition from static friction to sliding friction is obtained from the Eytelwein equation:

$$R = \frac{F_1}{F_2} = e^{\mu' \beta},$$

where $\mu' = \frac{\mu}{\sin\frac{\alpha_s}{2}}$.

As long as the belt-run forces are in this ratio, there will be no slip when power is transmitted. Belt run refers to that section of the belt between the belt inlet and outlet areas of two adjoining belt pulleys. A typical ratio for ribbed V-belts is $R = 4$ with a wrap angle of $\beta = 180°$.

To transmit the peripheral force

$$F_U = F_1 - F_2$$

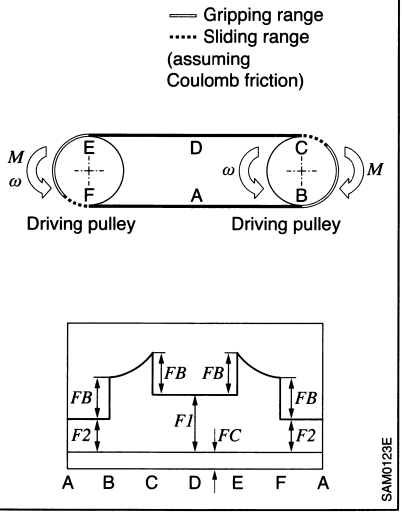
the pretensioning force F_{HL} ("hubload") is required (Figure 2). At high rotational speeds, the centrifugal force component F_c (centrifugal force) of the belt must also be taken into consideration.

Slip occurs on account of the change in tension within the belt during alternation between load and slack runs.

In spite of a constant transmitted torque, the belt tension decreases at the driving pulley and increases at the driven pulley (Figure 3). This variation in tension causes the belt to elongate over the respective pulley. In this elongation range, the belt loses its grip on the pulley and passes into a sliding state. The belt will slip if this sliding range stretches over the entire pulley wrap. In correctly configured drives the slip is under 2 % and the ribbed V-belts operate with an efficiency of more than 96 %. This configuration must be taken into account particularly in the case of belt-driven functions, such as engine starting and engine acceleration, since driving-half and undriven-half forces are reversed for this situation.

Figure 3: Strain distribution in the belt

F_1 Force on load side,
F_2 Force on slack side,
F_B Loading of belt by wrap,
F_C Loading of belt by centrifugal forces,
M Transmitted torque,
ω Angular frequency.



Design of ribbed V-belt

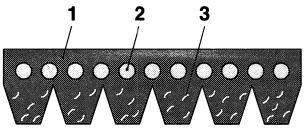
The ribbed V-belt is a composite of three components (Figure 4):
- fiber-reinforced rubber compound,
- tension members,
- backing or rubber coating.

The rubber compound forms the ribs and transmits the drive forces from the pulleys to the tension members. The main material used here is Ethylene Propylene Diene Monomer rubber (EPDM), and the rubber compound is fiber-filled for reinforcement.

The tension members accommodate the dynamic forces and transmit the drive power from the driving shaft (usually the crankshaft) to the accessories. They are usually made of nylon, polyester or aramide. These tension member materials

Figure 4: Structure of ribbed V-belt (cross-section)

1 Backing or rubber coating,
2 Tension members,
3 Rubber compound, fiber-filled.



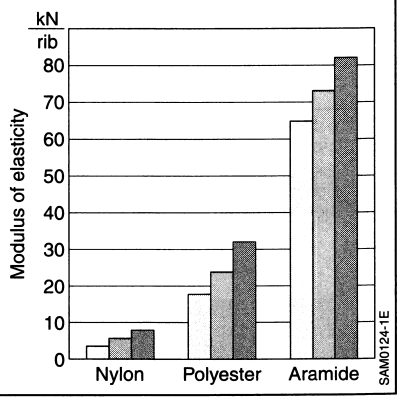
differ primarily in their very different moduli of elasticity (Figure 5).

The dynamic matching of the system can be optimized by selecting the appropriate tension member material. Tension-member materials with a high modulus of elasticity (e.g. aramide) are used in highly dynamic applications to influence the resonant frequencies of the belt drive. Depending on the choice of tension-member material, maximum dynamic belt forces of 390 N (for polyester) or up to 600 N (for aramide) per rib can be absorbed.

The back of the belt can be designed as a backing or as a rubber coating. The back of the belt forms a protective layer for the tension members. In most applications, it serves only to guide the belt quietly via the idler and tension pulleys. In some applications, however, it also serves to drive accessories which are subject to low load (e.g. water pump).

The rib profile can be manufactured by grinding, cutting or molding in a tool. The latter delivers the greatest precision in the rib geometry, helping to improve running time and noise generation. Fabric layers on the rib side additionally help to increase robustness.

Figure 5: Moduli of elasticity of different tension-member materials
The three columns for each material represent the spread.



Ribbed V-belt profile and belt pulleys

Usually the PK profile in accordance with ISO 9981 [1] is used in automobile applications (Table 2 and Figure 6).

The reference belt length L_B is determined on a two-pulley test rig under a defined pre-tension (ISO 9981). Here, the reference circumference U_B of the measurement belt pulleys is 300 mm (Figure 7). The reference belt length is then calculated according to

$$L_B = U_B + 2E\,.$$

Because the belt and pulley geometries in ISO 9981 are encumbered with a tol-

Table 2: PK belt profile and belt-pulley profile acc. to ISO 9981
Designation of a ribbed V-belt with six ribs of profile code PK and of reference length 800 mm: Ribbed V-belts 6PK 800.
Designation of corresponding ribbed V-belt pulley with reference diameter 90 mm: Ribbed V-belt pulley P6PK 90.

Dimensions in mm (Figure 6)	Belt	Profile grooves
Profile code	PK	K
Rib spacing s and groove spacing e	3.56	3.56
Perm. deviation tolerance for e		±0.05
Sum of perm. deviation tolerance for e		±0.30
Groove angle α_S		40° ±0.5°
Radius r_k at rib head Radius r_a at groove head	0.50	min. 0.25
Radius r_g at rib base Radius r_i at groove base	0.30	max. 0.50
Belt height h	4 to 6	
Test-pin nominal diameter d_s		2.50
2 h_s nominal dimension		0.99
2 δ		max. 1.68
Spacing f groove to thrust collar or ring		min. 2.5

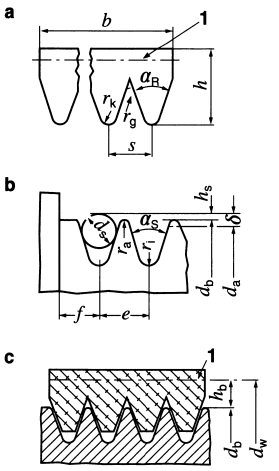
Belt width $b = n\,s$ (where n = number of ribs).
Effective diameter for profile K:
$d_w = d_b + 2\,h_b$, $h_b = 1.6$ mm.

erance band, it is absolutely essential for a detailed configuration to fall back on the characteristic values of the respective belt or pulley manufacturers.

The pulleys are made from steel, aluminum or plastic.

Figure 6: Belt profile

a) Belt cross-section,
b) Groove cross-section,
c) Determining effective diameter.
1 Position of tension member.
b Belt width,
h Belt height,
r_k Radius at rib head,
r_a Radius at groove head,
r_g Radius at rib base,
r_i Radius at groove base,
s Rib spacing,
e Groove spacing,
f Spacing groove to thrust collar or ring,
α_R Groove angle, belt,
α_S Groove angle, pulley,
h_s Allowance, test-pin nominal diameter to reference diameter,
d_s Test-pin nominal diameter,
d_a Outside diameter (diameter, groove-head profile),
d_b Reference diameter,
δ Allowance, test-pin nominal diameter to outside diameter,
d_W Effective diameter.



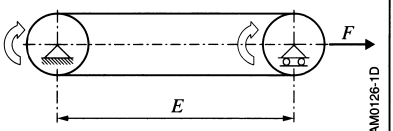
Drive system: accessory drive

The most important requirement of the accessory drive system is that it should drive all the accessories without slip. This must be guaranteed for all load states and under all ambient conditions over the entire service life of the engine. In modern engines with full drives – i.e. all the accessories are driven via an individual belt drive – maximum torques of up to 30 Nm (70 Nm for start/stop systems) and maximum outputs of 15to20 kW with all the accessories at full load are consequently transmitted via a five- or six-ribbed V-belt. The temperatures in the engine compartment are around –40 to +140 °C.

It is essential to utilize optimum system design to eliminate in particular slip noises, such as, for instance, the familiar V-belt squeaking in cold and damp weather conditions. It is also necessary to utilize design already to avoid belt noises caused by pulley misalignment. For belt-driven starting applications the shear loads to be transmitted by the ribs must most also be taken into account. Specially developed belt designs fulfill these requirements without damage caused by ribs shear off. In order to assess the service life, appropriate fatigue tests according to Wöhler must be carried out (see Wöhler fatigue tests).

Figure 7: Belt-length measurement acc. to ISO 9981

F Measurement force (pre-tension): 100 N per rib.
E Spacing of measurement belt pulleys, Reference circumference U_B of measurement belt pulleys: 300 mm.



Design criteria
Both manufacturer's own and cross-manufacturer computing programs are used to design accessory drives. Important input parameters are the arrangement of components, i.e. the drive configuration, the torque curves, the moments of inertia of the components, the torsional vibration of the crankshaft, and the belt data. With these data, it is possible to calculate and optimize the system geometry, such as, for example, belt-run lengths and wrap angle, the system natural frequencies, the slip limit angles, the belt-run forces, the belt-run vibrations, and also the service life of the belt.

Recommended minimum wrap angles

Crankshaft	150°
Alternator (I < 120 A)	120°
Alternator (I > 120 A)	160°
Power-steering pump, A/C compressor	90°
Tension pulley	60°

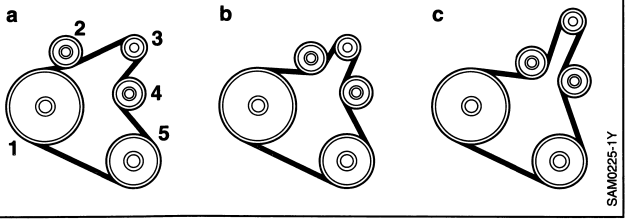
Misalignment and entry angle
In order to avoid unacceptable belt wear and noises, it is important that the entry angle of the belt into the grooved pulleys, depending on the belt design and the required service life, does not exceed 1.2 to 1.5°.

System natural frequency
The system natural frequency of the belt drive should not coincide in the idle range with the engine's ignition frequency as otherwise strong belt vibrations could occur on account of natural resonance.

Minimum diameters of pulleys and idler pulleys
In practice the smallest belt pulley is often located on the alternator in order to facilitate the high rotational speeds that are required there. Typical alternator pulleys have a diameter of 49 to 56 mm. Belt fatigue increases exponentially when very small pulleys are used; this must be taken into account when the belt is designed. Diameters not less than 60 mm are recommended for idler pulleys which generate strong belt-back deflection.

Belt-tensioning systems
Belt tension in accessory drives is today usually provided by way of automatic tension pulleys. These tension pulleys ensure that the belt is almost constantly pretensioned over the belt's service life by compensating belt elongation and belt wear. The design of the tension pulleys is significantly dictated by the space available. The system tensions are usually 40 to 70 N per rib.

Figure 8: Application examples for belt tensioner in start/stop systems
a) Two-belt-tensioner concept,
b) Pendulum belt tensioner variant 1,
c) Pendulum belt tensioner variant 2,
1 Crankshaft,
2 Tension pulley,
3 Starter-alternator unit,
4 Tension pulley,
5 Refrigerant compressor.



To monitor the dynamic response of the belt drive, common tensioning systems have a mechanical spring which works in combination with a damping unit.

Two tension pulleys are used in belt-driven start/stop systems. The tension pulleys either each have an independent spring element (two-belt-tensioner concept) or are connected to each other only by a torsion spring (pendulum belt tensioner). Start/stop belt drives with pendulum belt tensioners can be implemented with low belt pre-tension (Figure 8).

Stretch Fit drives

Elastic ribbed V-belts are also used occasionally in less complex belt drives. These use nylon as the tension-member material and do not require belt-tensioning systems. This belt type is pulled during fitting onto the pulleys by overstretching. The belt length must be designed in such a way that the belt pretensioning force is so high after fitting that, while wear and elongation over the required service life as well as the relevant ambient conditions are taken into account, sufficient pre-tension is still maintained without retensioning. Today, belt drives with up to four pulleys, i.e. up to three driven components, are possible.

Positive belt drives

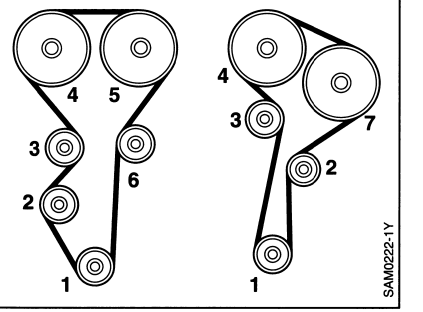
Application

Toothed belts in accordance with ISO 9010 [2] are used in timing drives to drive camshafts or the high-pressure fuel-injection pump synchronously with the crankshaft. The main advantages over competing means of drive – such as the gear drive or the chain – lie in the simplicity of the drive, the flexibility of the belt guidance, the low friction, the low noise emissions, and the capacity to compensate for high dynamic peak loads. Accessories such as oil or water pumps can be integrated into the drive or driven individually (Figure 9).

In modern applications, in contrast to earlier applications, it is possible in many cases to dispense with a replacement interval because of the use of innovative belt technologies and optimized design of the overall system. The low weight and the low costs also militate for the use of a toothed belt.

Figure 9: Application examples for toothed-belt drives

a) In-line gasoline engine,
b) In-line diesel engine.
1 Crankshaft,
2 Water pump,
3 Tension pulley,
4 Camshaft 1,
5 Camshaft 2,
6 Idler pulley,
7 High-pressure fuel-injection pump.

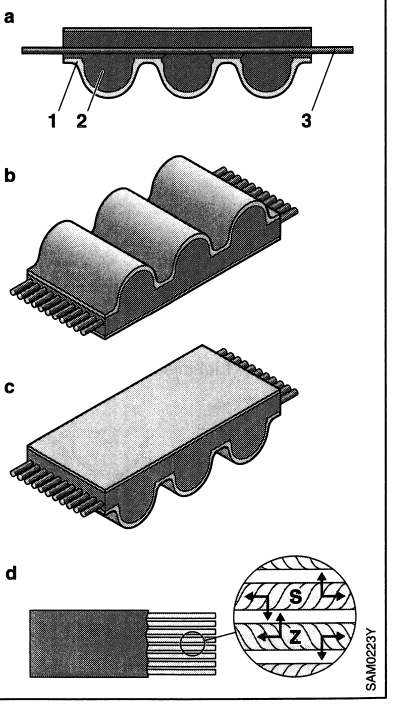


Forces and loads in the belt drive

In positive belt drives the torque is transmitted by the meshing of the toothed-belt teeth and a special toothed-belt pulley from the crankshaft to the toothed belt and from the toothed belt to the output components (e.g. camshaft, high-pressure fuel-injection pump or water pump). The actual power transmission is performed by the tension members embedded in the toothed belt. The toothing renders slip impossible, which means that toothed-belt drives can be used for synchronization.

Figure 10: Structure and characteristics of toothed belt

a) Longitudinal section,
b) Overhead view, tooth side,
c) Overhead view, back side,
d) Overhead view with depiction of tension members twisted in S- and Z-shapes.
1 Nylon fabric,
2 Rubber compound,
3 Tension member.



Structure of toothed belt

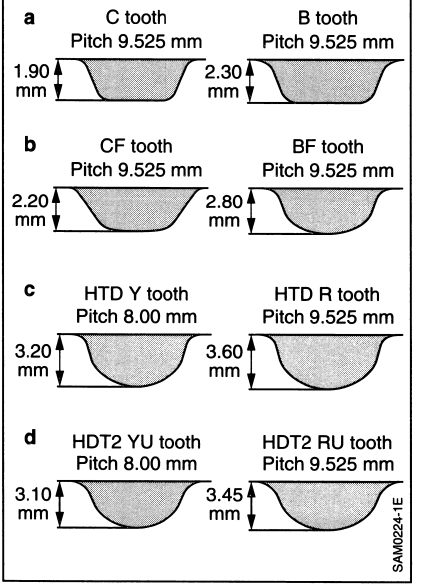
The toothed belt is a composite of three components (Figure 10):
- nylon fabric,
- rubber compound, and
- tension member.

The fabric is made of high-strength nylon and has an abrasion-proof and wear-resistant coating. It protects the rubber teeth against wear and from being shorn off, and improves the running performance.

The rubber compound consists of a high-strength polymer, and encloses the tension member on both sides. Polychloroprene (CR) was used in the original applications. The high demands placed on resistance to temperature and aging and dynamic strength means that today only Hydrogenated Nitrile Butadiene Rubber

Figure 11: Development of toothed-belt profiles

a) Power-Grip trapezoidal profile,
b) Power-Grip Power Function profile,
c) Power-Grip HTD profile,
d) Power-Grip HTD2 profile.



materials (HNBR) are still used in today's motor vehicles.

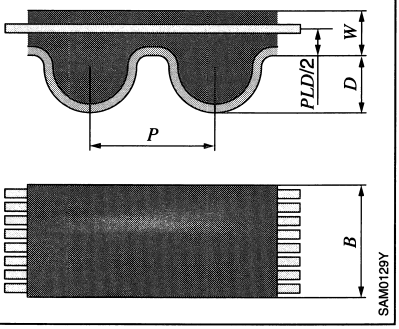
In some high-load applications belts with backing can also be used for increased reinforcement. This fabric reduces the risk of the belt running in the axial direction. It also improves belt service life in very cold ambient conditions (down to –40 °C), for example when the engine is started.

The tension member consists of twisted glass fibers – a material which is characterized by high tensile strength together with a high willingness to bend. As a result of the production process, the tension members are arranged in a spiral shape in the belt in the effective line, and twisted in pairs in S- and Z-shapes (Figure 10d). This ensures that the belt is subject to predominantly neutral running.

Toothed-belt profiles

The first camshaft belts were based on the classic Power Grip trapezoidal-tooth shape (Figure 11a), as was already known in industrial applications. On account of the increasing demands with regard to load transmission, jump protection and noise, circular-arc-like profiles (e.g. Power Grip HTD 2, High Torque Drive, Figure 11c and 11d) are almost exclusively used today. Compared with trapezoidal teeth, the forces with circular profiles are introduced more uniformly into the tooth, and stress concentrations are thereby avoided.

Figure 12: Characteristics, toothed belt
a) Longitudinal section, b) Overhead view.
P Pitch, *D* Addendum, *W* Rib thickness, *B* Width, *PLD*/2 Effective line distance (half the distance from outside diameter to effective line)



The pitch (Figure 12) is, in most cases, 9.525 mm for diesel engines, and 8.00 mm for gasoline engines. Higher forces can be transmitted with a larger pitch, while a smaller pitch has advantages with regard to noise and space.

Double-sided toothed belts can be used for applications where the direction of rotation is reversed (e.g. balancer-shaft drives).

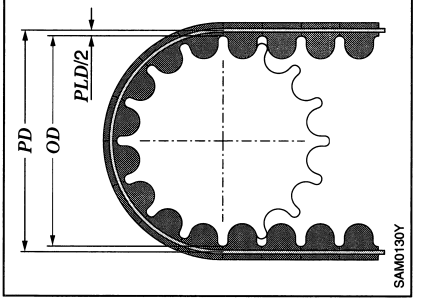
Toothed wheels

The associated toothed wheels are specified in ISO 9011 [3]. For the toothed wheel, the profile must be determined as a function of the diameter. The effective diameter *PD* is obtained from the number of teeth and the pitch, the outside diameter of the toothed wheel is accordingly reduced by the *PLD* (see characteristics in Figure 12 and Figure 13).

Drive system: toothed belt

The most important requirement of the toothed-belt drive system is that it should synchronize the camshaft in relation to the crankshaft over the service life of the engine. This is an important criterion for adhering to the consumption and emission values. It is possible to keep the elongation of the toothed belt under 0.1 % of the belt length by selecting the appropriate materials for the toothed belt, and by using an automatic tensioning system and optimized system dynamics. This results in 4-cylinder engines in a timing deviation of 1 to 1.5° referred to the crankshaft.

Figure 13: Characteristics, toothed wheel
PD Effective diameter, *OD* Outside diameter, *PLD*/2 Effective line distance.



The service-life requirements are currently in the range of 240,000 to 300,000 km, while the temperature requirements range from –40 °C in the cold state up to 150 °C engine-compartment temperature. Conspicuous noises caused for example by meshing are perceived to be offending and are therefore not acceptable. The efficiency of a toothed-belt drive is approx. 99 %. New materials and manufacturing processes today permit the use of toothed belts in an oil environment to drive oil pumps or to control the timing of camshafts. For the latter, it is possible to dispense with leading the camshaft out of the engine block and hence with the customary shaft seal on the main bearing. The advantages of this technology are the potential friction and noise reductions. The oil environment has not only a friction-reducing, but also a damping effect. This in turn has a positive impact, particularly for highly dynamic downsizing engines.

Design criteria

Both manufacturer's own and cross-manufacturer computing programs are used to design timing drives.

Important input parameters are the arrangement of components, i.e. the drive configuration, the torque curves of the components, and calculated from these the dynamic peripheral forces, as well as the belt data. With these data, it is possible to calculate and optimize the system geometry, such as, for example, belt-run lengths and wrap angles, and the system dynamics.

Recommended minimum wrap angles

Crankshaft	150°
Camshaft, high-pressure pump	100°
Accessory wheel	90°
Automatic tension pulley	50°
Idler pulley	30°
Damping pulley	10°

Periodic meshing

To avoid uneven belt wear, it is important to prevent the same belt teeth from engaging in the same wheel gaps. To assess this periodic meshing the tooth ratio of the number of teeth of the belt and the belt pulleys and the number of engine cylinders or the number of pistons in the high-pressure pump are considered.

Z	Tooth ratio
AM	Number of engine cylinders
AHD	Number of pistons in the high-pressure pump
ZR	Number of teeth on the belt
ZSNW	Number of teeth on the camshaft disk
ZSHD	Number of teeth on the high-pressure pump disk

$Z = ZR \cdot AM/ZSNW$
$Z = ZR \cdot AHD/ZSHD$

Periodic meshing does not occur when the following decimal values are avoided for the tooth ratio:

Z.0; Z.25; Z.33; Z.5; Z.66; Z.75.

Belt-run lengths
In order to avoid resonance noises at idle speed, it is important to ensure that free belt-run lengths are not in the area of 75 mm and 130 mm.

Minimum number of teeth and minimum diameters of toothed wheels and idler pulleys

Pitch 9.525 mm	17 teeth
Pitch 8.00 mm	17 teeth
Non-toothed idler pulleys	dia. 50 mm
Damping pulleys	dia. 28.5 mm

Tolerances of toothed wheels and idler pulleys
Concentricity, lateral running:
dia. ≤ 100 mm 0.1 mm
dia. > 100 mm 0.001 mm per mm dia.

Conicity of outside diameter
≤ 0.001 mm per mm wheel thickness.

Parallelism of bore to toothing
≤ 0.001 mm per mm wheel thickness.

Surface roughness
$R_a \leq 1.6$ µm.

Pitch error
- dia. ≤100 mm:
  ±0,03 mm gap/gap,
  max. 0,10 mm over 90°.
- dia. > 100 mm:
  ± 0.03 mm gap/gap,
  max. 0.13 mm over 90°.

Axial guidance
A toothed belt must be guided at least at one wheel on both sides with flanged wheels to prevent it running off the belt run. Essentially, it is important to ensure that toothed wheels with flanged wheels are exactly flush with the other wheels so that the belt is not deflected from its course. Limit values for belt-edge wear caused by excessively high run-off forces must be taken into consideration.

Belt-tensioning systems
The required constantly high belt tension, and compensation of the tension increase over temperature and belt elongation are usually provided in today's timing drives by tension pulleys. The design of the tension pulleys is significantly dictated by the space available. The most widely used system is the mechanical, friction-damped compact tensioner.

Essential design parameters for monitoring the system dynamics are the spring moment and the damping characteristic. Hydraulic tension pulleys are also used in some applications in the event of very high dynamic forces in the toothed-belt drive. These are characterized by their asymmetrical damping performance.

Tensioning systems can also be used in an oil environment if an appropriate adaptation of the damping level is taken into consideration.

References
[1] ISO 9981: Belt drives – Pulleys and V-ribbed belts for the automotive industry – PK profile: Dimensions.
[2] ISO 9010: Synchronous belt drives – Automotive belts.
[3] ISO 9011: Synchronous belt drives – Automotive pulleys.

Chain drives

Overview

Camshafts are used to control the gas exchange valves; these camshafts open and close the valves by means of direct actuation via bucket tappets or lever assemblies. Camshafts are driven on modern overhead valve engines, i.e. overhead camshafts, with the aid of a belt drive, while side-mounted camshafts are connected via a geartrain to the crankshaft pinion. Toothed belts, roller- or sleeve-type chains, and tooth-type chains are used to drive overhead camshafts. For fast-running assemblies – such as motor-racing engines for example – the drive is also effected with spur gears.

The most important criteria in deciding on the type of drive are costs, space, ease of maintenance, service life, and noise buildup. The great advantage of chain drives over toothed-belt drives is that they are completely maintenance-free over the entire service life of the engine, while toothed belts, depending on the application, have to be retensioned or changed according to the maintenance interval.

In addition to the camshaft, timing drives in modern engines frequently drive other assemblies, such as the oil pump and the high-pressure pump for gasoline direct injection or for common rail (Figure 1).

As both camshaft and crankshaft rotate nonuniformly and are exposed to torsional vibrations, the drive is subject to highly complex dynamic stresses. Furthermore, the torque demand of the high-pressure pump is subject to very strong periodic fluctuations, and thus generates further excitation of the timing-chain drive.

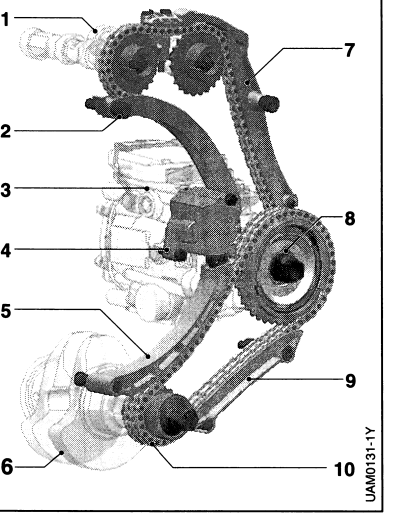
Timing chains

Designs of steel-link chains
In the case of standard chains, a distinction is made between roller- and sleeve-type chains. There are also simplex and duplex chains. The tooth-type chain is a special type of chain.

Roller-type chain
A steel-link chain is composed of inner and outer links. The inner link of a roller-type chain consists for example or two inner plates and two sleeves pressed into the plate eyes (Figure 2a). The outer link consists of two outer plates and two link pins which connect the inner and outer links. In a roller-type chain a roller is fitted over the sleeve.

The rotating rollers of a roller-type chain located over the sleeves roll with little friction on the tooth flanks of the sprocket

Figure 1: Timing-chain drive
1 Camshafts, 2 Tensioning rail,
3 Fuel-injection pump,
4 Chain-tensioner system, 5 Tensioning rail,
6 Crankshaft, 7 Guide rail,
8 Intermediate-shaft sprocket group,
9 Guide rail,
10 Crankshaft sprocket.



such that a different point on the circumference is always contacted. The lubricant between the rollers and sleeves helps to muffle noise and dampen shocks.

Sleeve-type chain
In the case of a sleeve-type chain (Figure 2b), the tooth flanks of the sprocket constantly touch the fixed sleeves at the same point such that the sleeves are subjected to an additional load. Perfect lubrication is therefore particularly important for these drives.

While having identical pitch – the chain pitch defines the distance from pin center to pin center – and force at break, sleeve-type chains have a larger joint surface than the corresponding roller chains, since the omission of the roller means that the pin diameter can be enlarged. A larger joint surface results in a lower joint-surface pressure and thus lower wear in the joints and a longer service life for the chain timing drive.

Sleeve-type chains have proven useful in camshaft drives subject to high loads in diesel engines, because increased wear resistance is called for here on account of the increased entry of soot into the engine oil.

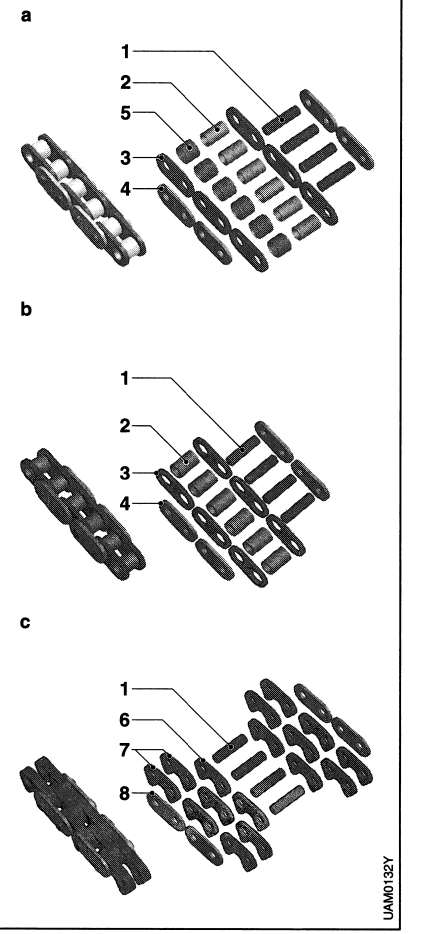
Tooth-type chain
A special design of steel-link chain is the tooth-type chain. On a tooth-type chain (Figure 2c), the plates are set out in such a way that they can transmit the force between the chain and the sprocket, whereas, in the case of roller- or sleeve-type chains, the connection with the sprocket is established via pin, sleeve, or roller.

Tooth-type chains can be built to virtually any desired width without the need for fundamental structural change. Guide plates, which are located either in the middle or on the outside (on both sides), are fitted to prevent the sprocket from running off.

A further variant of the tooth-type chain is the design that is toothed on both sides, which can then be used like a roller- and sleeve-type chain.

Sleeve-/tooth-type chain
Combining the advantages of a sleeve-type chain and a tooth-type chain creates a new type of chain for timing drives, a sleeve-/tooth-type chain (Figure 3). On the one hand, its special arrangement of tooth plates permits a smooth run into the sprocket; on the other hand, the use of sleeves reduces the wear on pins and

Figure 2: Chain designs
a) Roller-type chain,
b) Sleeve-type chain,
c) Tooth-type chain.
1 Link pin, 2 Sleeve, 3 Inner plate,
4 Outer plate, 5 Roller, 6 Tooth center plate,
7 Tooth inner plate, 8 Guide plate.



plates. This chain variant is recommended in all applications where excellent wear resistance and good acoustic and dynamic performance are important.

Selecting chain types

When choosing the chain types (construction and pitch), it is important, while adhering to the maximum sprocket diameter, to aim for numbers of teeth of more than 18 teeth in order to reduce dynamic influences by the polygon effect. The polygon effect occurs when a traction mechanism (chain) is positively driven by a drive wheel. With the sprocket at constant angular velocity the chain is accelerated and decelerated and consequently dynamically loaded.

Decades of experience have shown that some dimensions for roller-, sleeve- and tooth-type chains are particularly suitable for timing drives. These are 3/8" sleeve-type chains for diesel engines and 8 mm sleeve-type chains and 8 mm roller-type chains for gasoline engines. If in the development of a gasoline engine particular demands are made with regard to engine acoustics, 8 mm or 6.35 mm tooth-type chains should be used.

Requirements of chains

Four important factors determine the usage characteristics of timing chains: ultimate tensile strength, fatigue strength, resistance to wear, and acoustic behavior. A break may be caused by the static or dynamic load at break being exceeded. Especially in the case of timing drives, no uniform load will be encountered such that the fatigue strength of the chain is the strength-limiting quantity. The chain is subjected to a dynamic load as a result of the pulsating torques of the camshafts, the high-pressure pump, the rotational non-uniformity of the crankshaft, and the pulsating axial chain force caused by the polygon effect. Here, the safety limit to the fatigue strength of the chain must not be exceeded, since the number of such load alternations during an engine's service life is in each case greater than 10^8 load alternations (Figure 4).

In today's engines with precise timing and minimal clearance between piston and valve, minimal elongations due to chain wear can be achieved. Length increases caused by wear of just 0.2 to 0.5 % of the chain length for an engine mileage of up to 350,000 km are ensured today by the optimization of the pin/sleeve joint.

In particular direct-injection turbodiesel engines but also the latest models of direct-injection gasoline engines require timing chains that exhibit the highest levels of wear resistance.

To improve the wear resistance, the link pins are coated in a final operation with a layer of hard material, either chromium carbide or chromium nitride.

Figure 3: Sleeve-/tooth-type chain
1 Sleeve,
2 Tooth plate,
3 Guide plate,
4 Pin.

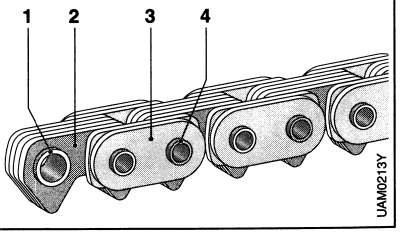
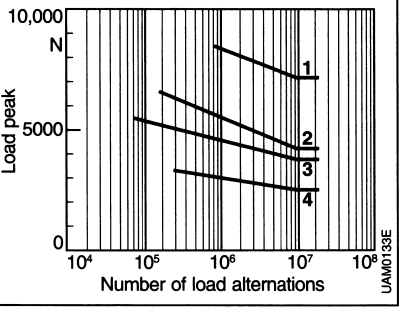


Figure 4: Fatigue-strength results for sleeve- and roller-type chains
1 Sleeve-type chain for DI applications,
2 Sleeve-type chain for prechamber diesel engines,
3 Roller-type chain for high-performance timing drives in gasoline engines,
4 Roller-type chain for oil-pump drives.



Design criteria for chain timing drives

A chain timing drive is with mass, rigidity, and damping an oscillatory system with several degrees of freedom. If subjected to corresponding excitation for example by the camshaft, crankshaft or high-pressure pump, this can, on account of interactions, cause resonance effects which lead to extreme timing-drive load.

A range of design criteria must be taken into account in the development of chain timing drives. These are:

- Chain length with an even number of links.
- Free run length (unguided chain section) if possible three to five chain links.
- Large radii at tensioning and guide rails for reducing friction.
- Inward tensioning direction of the chain tensioner.
- Floating mounting of the tensioning and guide rails.

Computational check of chain timing drives

Layout drawings are taken as the basis to create a chain-drive model which is used for the dynamic computation of timing drives. The timing drive is described by mass parameters, chain stiffness, and damping values.

Tensioning and guide rails are modeled as elastic bodies, the tensioning element is depicted with all the function elements. The simulation calculation delivers the necessary torsional-vibration angle deviations and all the force and moment characteristics.

Friction reduction of chain timing drives

To reduce the CO_2 emissions from an internal-combustion engine, every savings potential must be utilized to reduce fuel consumption. In this way, a friction-optimized timing-drive design can also make a contribution to reducing the CO_2 emissions.

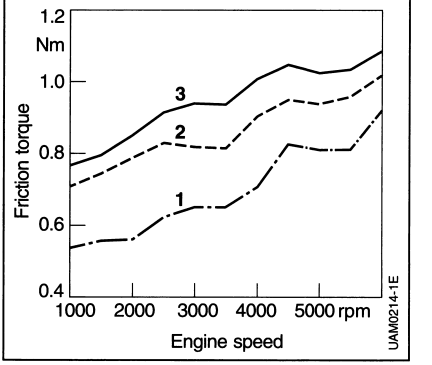
The decisive factor is the optimal layout of the chain line. Avoiding pronounced curvatures of the tensioning and guide rails reduces the normal forces and hence reduces the friction.

Chain-plate quality is of crucial importance to the frictional behavior of timing chains. The manufacture of chain plates involves the use of different blanking processes which create textures on the surface based on the manufacturing process used. Chains with precision-blanked plates deliver the best values.

Due to their internal friction and the contact of pins and plates, even timing chains of the most important types exhibit different frictional behavior. Figure 5 clearly shows the disadvantages of the tooth-type chain with identical overall width.

Figure 5: Friction torque of different chain designs

Chain pitch is the same for all designs at 8 mm.
1 Sleeve-type chain,
2 Roller-type chain,
3 Tooth-type chain.



Sprockets

The tooth shape of the sprockets is standardized for roller-, sleeve- and tooth-type chains (DIN ISO 606, [1]). DIN permits considerable degrees of freedom in the precision design of the sprocket toothing such that a tooth-space shape can be created between a minimum space and a maximum space. As a rule, sprockets with maximum tooth-space shape are used. This design permits, as a result of the low addendum and the larger tooth-space opening, the chain to enter and exit undisturbed even at higher chain speeds.

In fashioning the tooth shape of tooth-type chains each manufacturer applies its own design philosophy (e.g. noise- or wear-optimized) such that these chains do not conform to the standard. The design of the sprockets is dependent on the space conditions, the operating conditions, and the power transmission of the chain drive (Figure 6).

Sprockets made of carbon steel (e.g. Ck45) and alloyed steels (e.g. 16MnCr5) and wheels made of sintered materials (e.g. Sint D11) are used. Precision-blanked and machine-cut wheels with the heat treatment appropriate for the material are used.

Chain-tensioning and chain-guide elements

By using permanently acting clamping and guide elements (Figure 1) which are precisely matched to the engine in question, it is possible to optimize the chain drive in such a way that its service life corresponds to that of the engine.

Chain tensioner

The chain tensioner (Figure 7), mechanically or hydraulically acting, assumes a range of tasks in the timing drive. On the one hand, the timing chain is pre-tensioned in all operating conditions on the loose side under a defined load, even in the event of wear elongation that occurs during operation. Slack run refers to the non-pulled, as it were load-free chain section. A damping element – either friction or viscous damping – reduces vibrations to an acceptable level.

Mechanical chain tensioners without additional hydraulic damping are used, as

Figure 6: Sprockets
a) Sintered crankshaft gear,
b) Precision-blanked camshaft gear,
c) Sprocket with sealing surface,
d) Two-track sprocket for duplex chain.

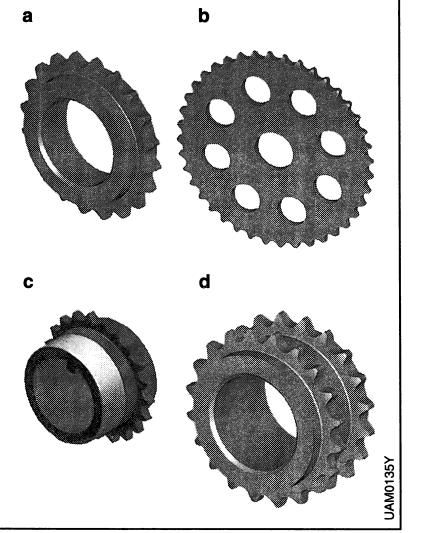


Figure 7: Types of chain tensioners
a) Mechanical chain tensioner,
b) Hydraulic chain tensioner.

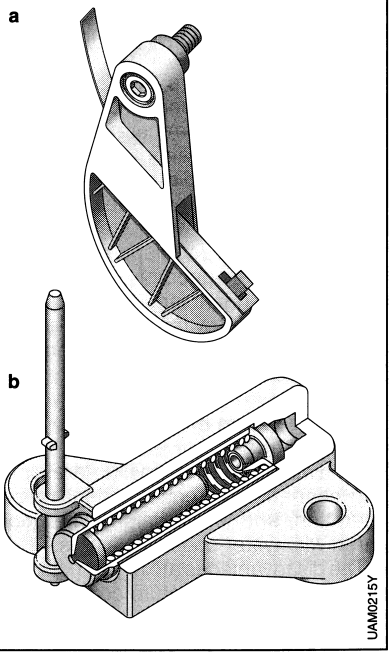
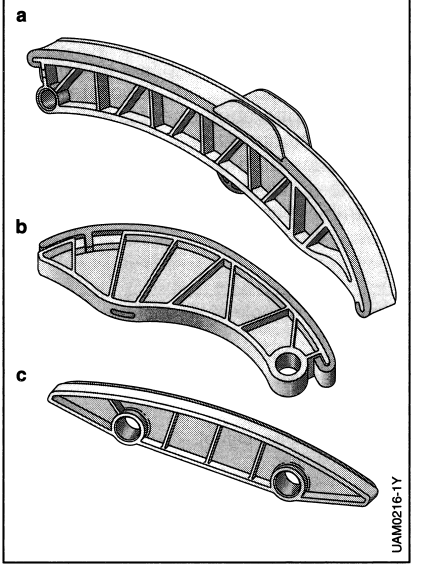


Figure 8: Types of guide and tensioning rails
a) Tensioning rail, two-component type,
b) Tensioning rail, two-component type, with lubricating duct,
c) Guide rail, one-component type.



a rule, in oil-pump drives which are subject to low load. In special cases, this mechanical chain tensioner can even be omitted completely.

Tensioning and guide rails

Simple rails made of plastic or metal (aluminum or sheet steel) with plastic covering and which are flat or bent - depending on the chain route - are sometimes used as tensioning and guide elements. In the case of newer types, the rails are for the most part inexpensively injection-molded from plastic and manufactured as a two-component design (base carrier and sliding lining).

Here, in the case of the tensioning rail, a friction lining of glass-fiber-free polyamide is sprayed or clipped onto a carrier of heat-resistant polyamide with 30 to 50 % glass-fiber content. The sliding rails are usually designed in plastic as one-component rails, and are used for chain guidance. Figure 8 shows some types of tensioning and guide rails.

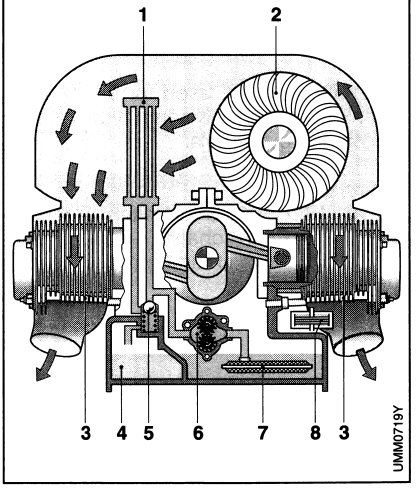
References

[1] DIN ISO 606: Short-pitch precision roller and bush chains, attachments and associated chain sprockets (ISO 606:2004).

Engine cooling

The process of converting any type of energy into mechanical energy to drive a vehicle creates heat as a waste product, whether it be by friction of different engine components against each other or by the combustion of fuel in the engine. In order to optimally utilize the conversion of the chemical energy of fuel into kinetic energy, this process must take place under controlled conditions and, to protect the engine and its components, also under controlled temperatures. The waste heat that is generated must be reliably dissipated to atmosphere under all operating and ambient conditions.

Figure 1: Principle of air cooling
1 Oil cooler,
2 Cooling-air fan,
3 Cylinder with cooling fins,
4 Oil pan,
5 Pressure-relief valve,
6 Oil pump (gear pump),
7 Lube-oil filter,
8 Thermostat.



Air cooling

Design and operating principle

Cooling air is routed by dynamic pressure or a fan around the external walls of the cylinder casing (Figure 1). These walls are finned in order to achieve better cooling results thanks to the increased surface area. The cooling-air quantity can be controlled for example by specifically restricting its flow at the vehicle's cooling-air inlets, which is achieved for example with thermostat-controlled flaps or by means of the rotational speed of a load- or temperature-dependent fan.

Heat absorbed by the engine oil is dissipated via cooling fins on the oil pan which are also located in the air stream.

Advantages and disadvantages of air cooling

The advantages of air cooling are simple and cost-effective construction, reliable operation, and lower weight when compared with water-cooling systems.

The disadvantages are higher noise generation, lower temperature consistency of all engine parts, and poorer thermal behavior at high specific power outputs.

Air-cooled engines struggle to comply with the currently required exhaust-emission limits.

Applications

Today, air cooling is mainly used for motorcycle and aircraft engines and in special applications. Extra-small engines in model-making applications also use air cooling.

The expected operating conditions for a vehicle can only dictate which cooling principle is chosen. In Arctic regions, where temperatures down to −50 °C are experienced, water-cooled engines cannot be used owing to the risk of the water freezing; for this reason air cooling must be used in these regions.

Water cooling

Design and operating principle

Cooling circuit

In contrast to air cooling, with water cooling the waste heat from the engine is not dissipated directly to the ambient air. Instead, the waste heat is absorbed by the coolant and transported via this medium to the main coolant radiator (Figure 2). Here, heat is exchanged between the coolant and the ambient air. The cooling-air flow rate is controlled by a fan, and the coolant is circulated in the circuit by a coolant pump.

Coolant cooling, also called water cooling, has become the standard in both passenger cars and commercial vehicles.

Coolant

Coolants are a mixture of water, antifreeze (mostly ethylene glycol), and corrosion inhibitors. The inhibitors prevent the components through which coolant flows from corroding (see Coolants). Adding antifreeze in concentrations of 30 to 50 % firstly stops the coolant from freezing even at temperatures down to – 25 °C and secondly raises the coolant boiling temperature to allow coolant temperatures of up to 120 °C at a pressure of 1.4 bar in passenger cars.

Radiator designs

The cores of the coolant radiators in modern passenger cars are almost exclusively made of aluminum. Aluminum radiators are also being used to an ever-increasing extent in a wide range of commercial vehicles and trucks throughout the world. There are two assembly variants: brazed radiators and mechanically joined or assembled radiators.

Flat-tube/corrugated-fin systems

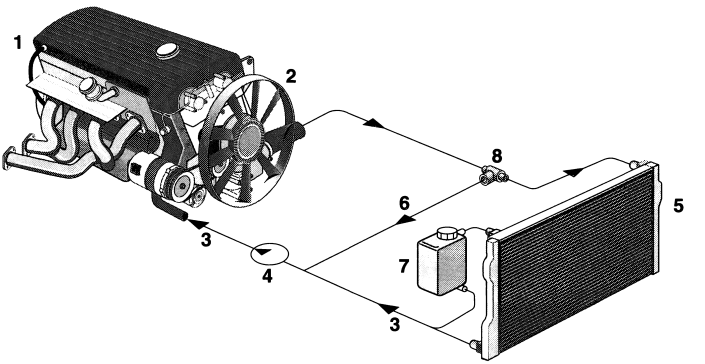
For cooling high-performance engines, or when space is limited, the best solution is a brazed flat-tube and corrugated-fin radiator layout with minimized aerodynamic resistance on the air-intake side.

Construction

The radiator core consists of series of flat tubes with corrugated fins located in between (Figure 3). The flat tubes lead into the openings in the radiator base. During the manufacturing process the tips of the corrugated fins are soldered to the

Figure 2: Passenger-car cooling system

1 Engine,
2 Fan,
3 Coolant line (arrow indicates direction of flow),
4 Coolant pump,
5 Main coolant radiator (flat-tube/corrugated-fin system),
6 Bypass line (secondary cooling circuit),
7 Expansion tank with pressure-relief valve (integrated in cap),
8 Thermostat.



flat tubes and the flat tubes to the radiator base in an oven. The solder material is applied as a layer to the materials.

The radiator tank ensures that the coolant is distributed throughout the core. These tanks are made of fiberglass-reinforced polyamides, and are injection-molded with all connections and mountings in a single unit. They are flange-mounted to the radiator core and sealed by an integrated elastomer sealing element.

Operating principle

The coolant is routed through the radiator tank to the radiator core. It flows through the flat tubes before emerging again from the opposite radiator tank. The hot coolant in the flat tubes dissipates heat to the corrugated fins, while the cooling air flowing through the radiator core cools the corrugated fins. The soldered connections guarantee a good heat transfer from the tubes to the fins. Gills in the corrugated fins, which are brought together in gill fields, bring about additional swirling of the cooling air. This results in improved heat dissipation.

Radiator tanks are designed in such a way as to distribute the coolant as uniformly as possible to all the flat tubes. The tubes have minimal depth (approx. 2 mm with wall thicknesses of 0.2 to 0.3 mm) so that the full coolant flows past the tube walls as closely as possible, thereby ensuring a good heat transfer to the fins.

Tube/fin systems

The less expensive, mechanically assembled tube/fin system is generally employed for applications with less powerful engines or when more space is available.

Construction

When the radiator is assembled mechanically, the cooling grid is formed by mounting stamped cooling fins around round, oval or flat-oval tubes (Figure 4). The distance between the cooling fins is maintained by trapezoidal cutouts. In the manufacturing process the tubes are

Figure 3: Flat-tube/corrugated-fin system
1 Radiator tank,
2 Radiator base,
3 Gasket,
4 Flat tube,
5 Corrugated fins,
6 Gill zones.

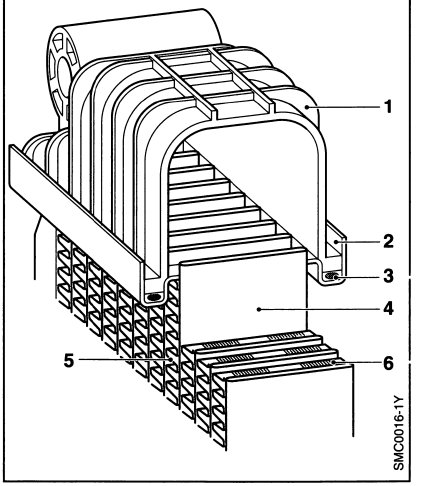
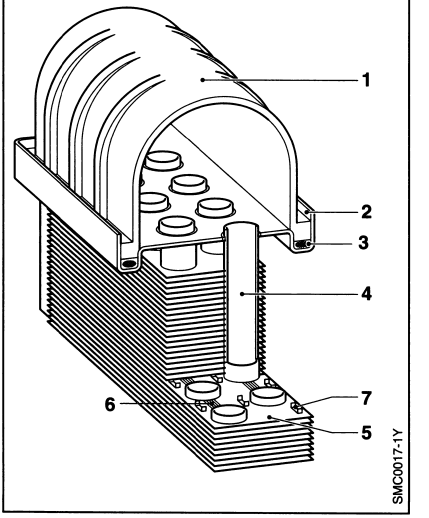


Figure 4: Tube/fin system
1 Radiator tank,
2 Radiator base,
3 Gasket,
4 Round tube,
5 Cooling fin,
6 Gill zones,
7 Turbulators.



initially inserted loosely into the cooling fins and flared with a tool. This creates an internal connection between the tubes and the fins. The connection is established exclusively by the mechanical flaring of the tube. However, the heat transfer between the tubes and the fins is much worse than in the soldered systems.

Operating principle
The cooling air flows through the radiator core between the cooling fins. These are corrugated or slotted at right angles to the direction of air flow. They are – like the corrugated fins in soldered radiators – also fitted with turbulators and gill zones, inducing air turbulence and improving the cooling effect.

Options for increasing radiator performance
Gills and a larger number of corrugations are integrated into the cooling fins on the cooling-air side to enhance thermal transfer. The gills induce turbulence in the cooling air flowing through, and the larger number of corrugations produces a larger surface area. These measures increase the quantity of heat that can be dissipated.

Further measures to increase cooling efficiency include the use of tubes with the lowest possible width and wall thickness, and turbulators on the coolant side, provided the attendant pressure losses remain within acceptable limits. The negative effect of these performance-enhancing measures is the higher pressure drop on both the cooling-air and coolant sides, which must be compensated by a higher-capacity fan and on the coolant side by a more powerful coolant pump.

Coolant expansion tank
Mounted at the highest point of the cooling circuit is the airtight expansion tank; this tank has a filler neck through which coolant is poured into the cooling system. When the engine is running the expansion tank takes up excess coolant that expands due to heating. Where necessary, it returns it to the cooling circuit. As the coolant heats up, so too the system pressure in the cooling system increases. This causes the coolant boiling point to rise.

The expansion tank's air volume must be large enough to absorb the coolant's thermal expansion during rapid pressure buildup, and prevent the coolant from boiling over under the permitted operating conditions. At excessively high operating temperatures and thus also excessively high operating pressures the expansion tank protects the cooling circuit by means of a pressure-relief valve.

The expansion tank also provides a reliable escape channel for pressurized gases, preventing cavitation of the kind that tends to occur on the intake side of the coolant pump.

Expansion tanks are injection-molded in plastic (generally polypropylene), although simple designs can also be blow-molded to shape. A system of hoses is normally used to connect the expansion tank to the cooling system. The expansion tank is mounted in the engine compartment at the location representing the highest point in the cooling system to ensure air is expelled effectively. In some cases, the expansion tank can form a single unit with the radiator tank, or the two can be joined by means of a flange or plug connection.

The position and shape of the filler neck can be used to limit capacity, thus preventing overfilling. An electronic level sensor can be fitted to monitor the level of the coolant. The coolant level can also be monitored by manufacturing the expansion tank completely or partially in natural-color, transparent plastic with molded level markers. However, colorless polypropylene is sensitive to ultraviolet rays. For this reason, the transparent part of the expansion tank should not be exposed to direct sunlight.

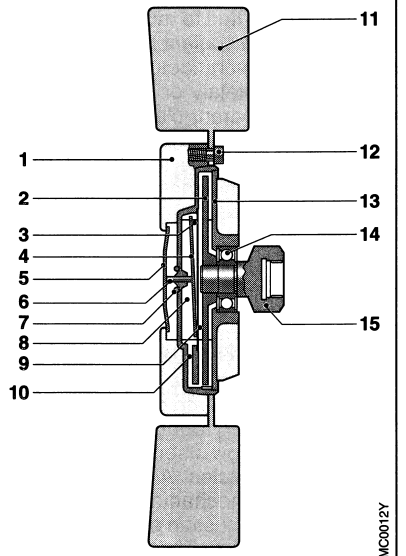
Cooling-air fan
Design
Since motor vehicles also require substantial cooling capacity at low speeds, force-air ventilation is required for the radiator. Injection-molded plastic fans with drive-power ratings extending up to 30 kW are now used as standard in commercial vehicles. A fan of this type is driven by a mechanical coupling with the internal-combustion engine, for example via a belt drive. The fan can also be mounted directly on the crankshaft.

Single-piece injection-molded plastic fans are generally used in passengers cars; these are usually driven by DC brush motors or brushless DC motors. These motors are installed in the fan hub. The electrical drive power is around 400 W in the small-car segment, and extends up to 1 kW in the luxury-class segment and in offroad vehicles. Although blade design and arrangement can be selected to provide relatively quiet running, these fans have considerable noise levels at high rotation speeds.

In some passenger-car applications, in particular offroad vehicles with very high engine power outputs in combination with equipment options for operation in hot climates and equipment variants for diesel engines and air-conditioning systems, the capacity of the electric drive is no longer sufficient to provide the volume of air necessary to ensure effective cooling. Fan-output figures in excess of 1 kW can only be achieved when the fan is driven mechanically by the engine. However, this is only possible with a longitudinally installed engine.

Figure 5: Bimetall-controlled viscous clutch

1 Cover with cooling fins,
2 Primary disk,
3 Valve bore,
4 Valve lever,
5 Thermal bimetallic strip,
6 Operating pin,
7 Seal,
8 Supply chamber,
9 Working chamber,
10 Return bore,
11 Fan,
12 Mounting bolt,
13 Basic body,
14 Ball bearing,
15 Flanged shaft.



Control of electrically driven fans

Depending on vehicle and operating conditions, the unassisted air stream can provide sufficient cooling up to 95 % of the time. It is thus possible to economize on energy that would otherwise have to provide the power to drive the fan. For this purpose, electric fans use a multistage or continuous control system that specifically adapts the operating periods and speed of the fan to the required cooling capacity. A multistage control system may consist of relays and series resistors, while continuous variable control necessitates the use of power electronics. Electric thermostatic switches or the engine control unit supply the input signals for the control system.

Drive of mechanical fans

The fluid-friction clutch (viscous clutch) is a mechanical-drive arrangement of proven effectiveness for use in commercial vehicles. It essentially consists of the following assemblies (Figure 5):

- Input or primary section (flanged shaft and primary disk)
- Output or secondary section (cover, basic body)
- Control facility for controlling and regulating the oil filling (working fluid)

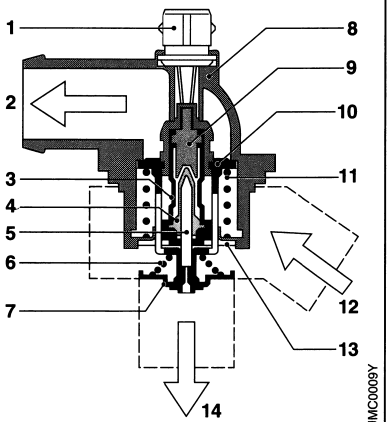
The secondary section of the viscous clutch is divided into a working chamber and a supply chamber. When the viscous clutch is not cut in, only a small amount of working fluid is in the working chamber, such that only a very small amount of torque can be transmitted due to the high degree of slip between primary disk and secondary section. The viscous working fluid is a silicone oil.

As the temperature increases, the bulging of the bimetallic strip causes the operating pin to open the valve. Silicone oil flows from the supply chamber into the working chamber. This reduces the slip between primary disk and secondary section – in turn increasing the fan speed. The fan speed and thus the cooling capacity are infinitely increased.

As the temperature decreases, the bimetallic strip cools and causes the operating pin to close the valve slowly. The silicone oil flows through the pump body back into the supply chamber. The fan speed decreases. The amount of silicone oil located in the working chamber determines the output power transmitted by the clutch and therefore the speed of the fan.

A distinction is made between types of viscous clutch, based on how the valve is actuated: Firstly, as already described, the temperature-dependent, self-governing clutch which varies its speed infinitely by means of a bimetallic element, an operating pin, and a valve lever. The controlled variable is the temperature of the air leaving the radiator, and thus indirectly the temperature of the coolant. The electrically activated clutch is now used in 95% of heavy-duty commercial vehicles. In this clutch type the valve lever is electromagnetically actuated and thus the oil quantity in the working chamber is regulated. Instead of just one controlled variable, a wide range of input variables is used for control purposes. These are usually the temperature limits of the various cooling media.

Figure 6: Map-controlled thermostat
1 Connector, 2 Connection to radiator,
3 Housing of working element,
4 Elastomer insert,
5 Plunger,
6 Bypass spring,
7 Bypass valve,
8 Housing,
9 Heating resistor,
10 Double-acting disk valve (main valve),
11 Main disc spring,
12 Connection from engine,
13 Tie bar,
14 Connection to engine (bypass).



Regulation of coolant temperature

A motor vehicle's engine operates in a very wide range of climatic conditions and with major fluctuations in engine load. The consequences of this are significant variations in coolant and engine temperatures, which in turn result in increased engine wear, unfavorable exhaust-gas composition, higher fuel consumption, and unreliable vehicle heating. The temperature of the coolant must be regulated in order to counteract these unwanted side effects and to keep the coolant and engine temperature as constant as possible.

Expansion-element-regulated thermostat

A temperature-sensitive thermostat incorporating an expansion element proves to be a robust regulator operating independently of the changing pressure conditions in the cooling system. It is very similar in design to the map-controlled thermostat (Figure 6), but does not have a heating resistor. The expansion element used in this type of thermostat actuates a double-acting disk valve (main valve), which – until the operating temperature is reached – closes the connection to the radiator while simultaneously releasing coolant flow from the engine outlet to the bypass line (see Figure 2). The coolant flows uncooled back into the engine ("secondary cooling circuit").

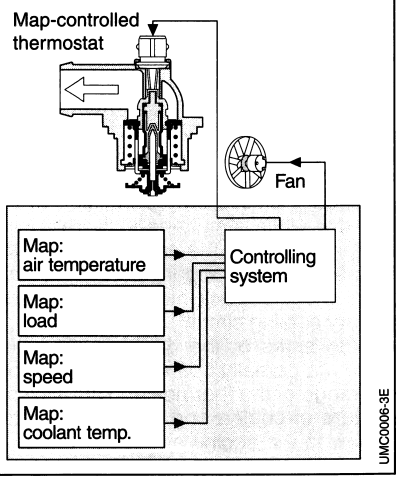
Both sides of the double-acting disk valve are partially opened within the control range of the thermostat. This allows a mixture of cooled and uncooled coolant to flow to the engine at such a rate as to maintain a constant operating tempera-

ture. At wide-open throttle (full load) the opening to the radiator is fully opened and the bypass line closed off ("primary cooling circuit").

Electronic map-controlled thermostat
Further possibilities to regulate the coolant temperature are permitted when a map-controlled thermostat is used. An electronically controlled thermostat differs from purely expansion-element-regulated thermostats in that the opening temperature can be controlled. A map-controlled thermostat features a heating resistor that is used to additionally heat the expansion element (Figure 6). This has the purpose of enlarge the opening of the double-acting disk valve to the radiator, thereby reducing the coolant temperature. The heating resistor is controlled by the engine management system in order to ensure that the engine operating temperature is optimally adapted to the operating conditions. The information necessary for this purpose is stored in the form of program maps in the engine-management system (Figure 7).

Raising the operating temperature in the part-load range and reducing the operating temperature at wide-open throttle provide the following benefits:

Figure 7: Electronic control of coolant temperature



- Lower fuel consumption,
- Low-pollutant exhaust-gas composition,
- Reduce engine wear,
- Improved heating efficiency of the vehicle interior.

Radiator design
The size and therefore the cooling capacity of a specific radiator can be determined by calculation, based on correlation equations derived from tests relating to thermal transfer and flow-pressure loss. The mass of air that flows through the radiator in the vehicle is a decisive factor and depends on the driving speed, the resistance to flow in the engine compartment, the resistance to flow of the radiator, and the efficiency of the fan.

The primary objective of radiator design is to maintain the coolant temperature at the engine outlet below a maximum permissible value under given operating conditions. Because at low driving speeds the dynamic pressure of the cooling air and thus the cooling-air flow rate in the radiator are very low, the cooling-air flow required for adequate cooling must be ensured either by a high-output fan or by a radiator with low flow resistance, while a high air-mass flow rate facilitates the use of smaller radiators with high flow resistance. However, the latter causes high energy consumption if the air-mass flow is generated by a powerful fan.

The task of determining the most favorable solution in terms of technical feasibility and economic efficiency is a matter of optimization that is best solved by the application of simulation tools. The most suitable and effective simulation tools describe all components that have an influence on air-mass flow. They depict the radiator as an integral heat-transfer medium or heat exchanger. The simulation results are checked by way of on-vehicle tests conducted in wind tunnels.

Intercooling (charge-air cooling)

Trends in engine development show a constant increase in specific engine output power. This development goes hand in hand with the current transition from naturally aspirated engines to supercharged engines and ultimately to supercharged, intercooled engines. The need for intercooling (charge-air cooling) is attributed to the higher air density levels associated with supercharging systems and therefore the amount of oxygen available in the combustion air. Intercooling also reduces exhaust-gas emissions from supercharged diesel engines. If there was no intercooling (charge-air cooling) employed on supercharged spark-ignition (SI) engines, appropriate steps would need to be taken to prevent engine knocking attributed to mixture enrichment or retarded ignition timing. Consequently, intercooling indirectly serves to reduce fuel consumption and exhaust emissions.

Design variations

Basically speaking, both the ambient air and the engine coolant can be employed to cool the charge air. With only few exceptions, air-cooled intercoolers are now used both on passenger cars and commercial vehicles.

Air-cooled intercoolers

Air-cooled intercoolers can be mounted in front of or next to the engine radiator, or even completely separate from the adiator at a different location in the engine compartment. A separately positioned intercooler can utilize either the unassisted vehicle air stream or its own fan. With the intercooler located in front of the engine radiator, the cooling-air fan ensures sufficient air flow even at low vehicle speeds. However, a drawback of this arrangement is that the cooling air is itself heated in the process. To compensate for this effect, the higher temperature of the inflowing cooling air must be taken into account in the design of the engine radiator.

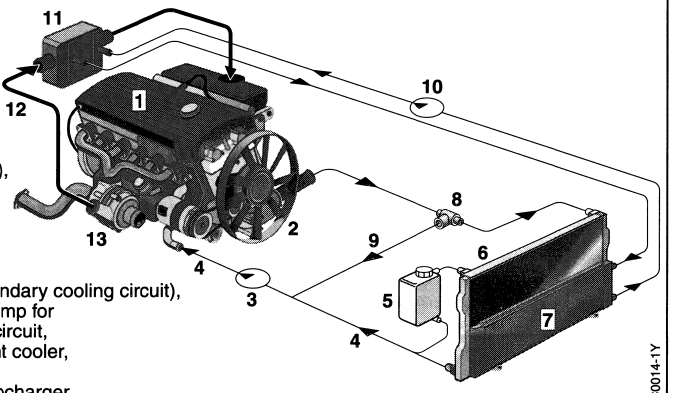
The system of corrugated aluminum fins and tubes employed for the intercooler core is similar to the design of the engine coolant radiator. Wide tubes with internal fins provide superior performance and structural integrity in actual practice. The fin density on the cooling-air side is relatively low and corresponds approximately to the density of the inner fins in order to achieve a good distribution of thermal-transfer resistance.

Wherever possible, the plenum chamber is injection-molded in fiberglass-reinforced polyamide as a single casting incorporating all connections and mounts. The plenum chambers which are subject

Figure 8: Passenger-car cooling system with intercooling

Engine cooling and indirect intercooling with separate low-temperature circuit.

1 Engine,
2 Fan,
3 Coolant pump,
4 Coolant line (arrow indicates direction of flow),
5 Expansion tank with pressure-relief valve (integrated in cap),
6 Main coolant radiator,
7 Low-temperature coolant radiator,
8 Thermostat,
9 Bypass line (secondary cooling circuit),
10 Auxiliary water pump for low-temperature circuit,
11 Charge-air coolant cooler,
12 Charge-air line,
13 Exhaust-gas turbocharger.



to increased stresses on the charge-air inlet side are injection-molded from highly heat-resistant PPA (polyphthalamide) or PPS (polyphenylene sulfide). They are flange-mounted to the radiator core and sealed by an integrated elastomer sealing element. Plenum chambers which feature undercut shapes or are intended for high-temperature applications are cast in aluminum and welded to the core.

Coolant-cooled intercoolers
Coolant-cooled intercoolers can be installed in virtually any location in the engine compartment, as there are no technical difficulties to supplying the system with coolant. In addition, thanks to its modest dimensions, this type of intercooler requires substantially less space than the air-cooled intercooler. Coolant-cooled intercoolers have a high power density. However, coolant at a very low temperature must be available in order to effectively cool the charge air. This requirement is of particular significance in commercial vehicles and heavy-duty trucks as, in this case, it is necessary to heat the charge air to a level of 15 K above the ambient temperature. Because this stipulation cannot be implemented with the temperature level of the normal cooling circuit at approx. 100 °C, it is necessary, in order to cool the charge air, to install a low-temperature radiator to ensure coolant is available in a separate circuit at the required temperature level.

Diffusion rate, intercooler
The diffusion rate Φ of an intercooler is particularly important in assessing intercooler performance. It defines the relationship between charge-air cooling efficiency and the charge-air/cooling-air temperature differential:

$$\Phi = \frac{T_{1E} - T_{1A}}{T_{1E} - T_{2E}}.$$

The equation's elements are:
T_{1E} Charge-air inlet temperature
T_{1A} Charge-air outlet temperature
T_{2E} Cooling-air or coolant inlet temperature

For passenger cars: $\Phi = 0.4...0.7$
For commercial vehicles: $\Phi = 0.9 - 0.95$

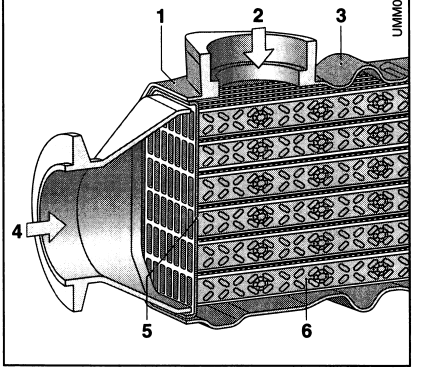
Exhaust-gas cooling

Exhaust-gas recirculation is used in diesel engines to reduce raw nitrogen-oxide emissions (NO_x). In gasoline engines it serves both to protect components and to dethrottle in the part-load range by charge dilution (see Exhaust-gas recirculation). Exhaust-gas recirculation delivers the best results when the recirculated exhaust gas is cooled in an EGR cooler.

Design of an EGR cooler
A central aspect in the design of an EGR cooler is, as well as its output, its durability. The former is achieved either by specially shaped tubes (winglets) which generate a high level of turbulence in the tube or by internal fins in the tubes which increase the surface area for heat exchange. The essential characteristic here is the tendency of an EGR cooler towards fouling. This can be significantly influenced by a suitable design of the winglets or tube fins. Figure 9 shows such an EGR cooler in longitudinal section. The winglet tubes direct the hot exhaust gas and are themselves surrounded by coolant.

It must be borne in mind with regard to durability that the hot exhaust gas may

Figure 9: EGR cooler
1 Stainless-steel housing (manufactured by internal high-pressure forming),
2 Coolant inlet,
3 Length compensation,
4 Exhaust-gas inlet,
5 Stainless-steel cooler base,
6 Stainless-steel winglet tubes.



condense in the cooler. The condensate produced has a lower pH value and attacks the material like a strong acid. For this reason, the gas-directing parts of EGR coolers are frequently made of stainless steel.

When the EGR cooler is in operation the tubes get hotter than the housing. Suitable design measures must be used to compensate for the linear difference in the cooler.

The issue of vibration also plays an important role in view of the fact that the cooler is usually bolted directly on the engine. The EGR cooler must therefore be as rigid as possible in its design. The first natural frequency of an EGR cooler should at least be above the first natural frequency of the engine block. For this reason, the brackets of such EGR coolers are specifically very rigid and sometimes multiply ribbed in design.

The cooler housing is often made of stainless steel, which can be strengthened further by cold-working. The exhaust pipes connected to the cooler should be isolated at the engine end. This isolation is achieved by bellows or by a comparable structure.

Cooling circuit

The coolant comes primarily from the engine cooling circuit, with the EGR cooler being cooled by coolant flowing out of the engine. A variant of this arrangement is "two-stage exhaust-gas cooling". Here the exhaust gas is cooled in a first stage as described above. There is also a second cooling stage, which is supplied by a cooling circuit independent from the engine (low-temperature circuit). Because the temperatures in this second cooling stage are much lower (approx. 10 to 20 K above ambient temperature), the exhaust gas can also be cooled significantly below the otherwise normal values, increasing the effect of reducing nitrogen oxide. The two cooling stages are either integrated together in a single housing or designed separately as two independent components.

Oil and fuel cooling

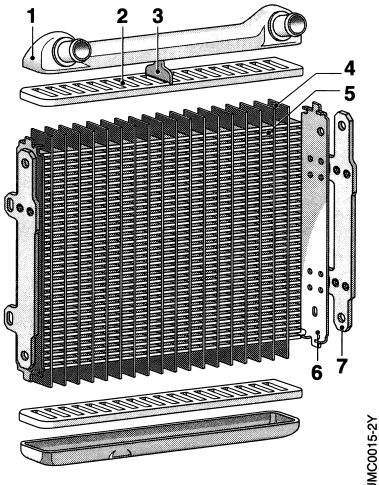
Oil coolers are often needed in motor vehicles to cool both the engine oil and transmission oil. They are used when the heat losses from the engine or transmission can no longer be dissipated via the surface of the oil pan or the transmission with the result that the permitted oil temperatures are exceeded.

Air- or coolant-cooled oil coolers are used to suit their application profile.

Oil-to-air coolers

Oil-to-air coolers are predominantly made of aluminum. In most cases, soldered flat-tube/corrugated-fin systems (Figure 10, similar to engine radiators) with a high power density are used. Mechanically assembled systems with round tubes and fins are less commonly used. Turbulence inserts are brazed in the flat-tube system to increase cooling capacity and strength (to resist high internal pressures).

Figure 10: Oil-to-air cooler (flat-tube/corrugated-fin system)
1 Collector tank with oil connections,
2 Base,
3 Partition,
4 Flat tube with turbulence insert,
5 Corrugated fin,
6 Side section,
7 Bracket.



The principle of oil-to-air cooling is ideal for use in commercial vehicles and powerful passenger cars to cool the transmission oil. To ensure good ventilation, they are ideally mounted upstream of the coolant radiator, but can also be mounted elsewhere in the engine compartment. Oil-to-air coolers which are not located upstream of the main coolant radiator and are thus outside the range of action of the main fan must be supplied with cooling air accordingly, for example by exposure to dynamic air pressure or by a separate electric fan.

Oil-to-coolant coolers

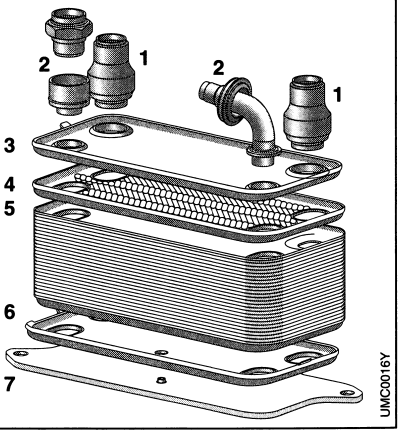
Aluminum stack designs employed in oil-to-coolant coolers have largely replaced stainless-steel disk coolers, double-tube oil coolers, and aluminum forked-tube coolers.

Disk-stack oil coolers

Disk-stack oil coolers are made of individual disks with turbulence inserts inserted between the disks (Figure 11). The upright edges of the disks fit together in a casing. Passages connect the channels formed by the disks in such a way that coolant and oil flow through alternate channels.

Figure 10: Oil-to-coolant cooler (disk-stack oil cooler)

1 Oil connections,
2 Coolant connections,
3 Cover,
4 Disk stack (oil passage),
5 Disk stacks (coolant passage),
6 Reinforcement plate,
7 Base plate.



Double-tube oil coolers and flat-tube oil coolers

These cooler types are mounted directly in the radiator tank on the outlet side of the main radiator (coolant radiator). The radiator tank thus assumes the function of the coolant-side housing for the oil coolers. The double tube of the oil cooler is formed by an outer tube and an inner tube inside it with a turbulence insert in between. The two tubes are soldered to each other at the ends. The transmission oil passes through the space between the inner and outer tubes, around or through which already cooled coolant flows.

Double-tube coolers are used as transmission-oil coolers in passenger cars and commercial vehicle in the lower power segment up to approx. 2.5 kW. As power requirements increase up to 4 kW the double tube is replaced by numerous flat tubes connected to each other by turbulence plates on the coolant side. The flat tubes are interconnected by openings at their ends. The flat tubes are also provided on the oil side with turbulence plates, which are soldered to increase both cooling capacity and strength.

Disk oil coolers

Disk oil coolers are mounted between the engine block and the oil filter. They have a separate casing and a central channel for the oil to pass through. The oil flowing back from the oil filter is routed through a labyrinth of perforated disks separated by turbulence inserts. This labyrinth is cooled by coolant flowing from the main circuit through the casing.

Forked-tube coolers

Forked-tube coolers are made of finned forked tubes through which the coolant flows. On the oil side, they have no casing and must therefore be integrated into the oil-filter housing or the oil pan.

Engine-oil coolers for commercial-vehicle applications

The engine oil in commercial vehicles is generally cooled by stainless-steel disk-stack coolers or aluminum flat-tube cool-

ers without a casing on the coolant side. They are accommodated in an extended coolant duct in the engine block.

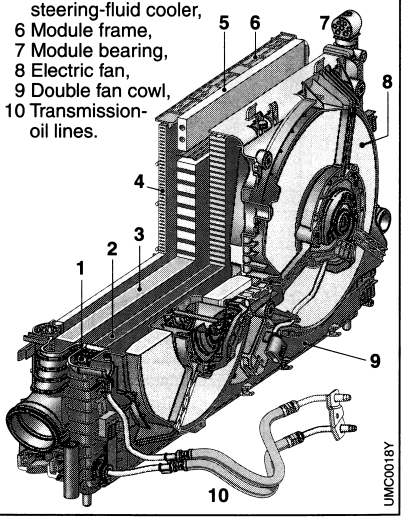
Fuel coolers

Fuel coolers are installed in modern diesel engines in order to cool excess diesel fuel down to permissible levels. This excess diesel fuel heats up during the injection process as the result of compression in the high-pressure pump, before it is routed back via the return line to the fuel tank. The excess diesel fuel that is returned without cooling to the tank has a temperature of well over 70 °C and heats the diesel fuel in the tank. If now the tank were almost empty, the temperature of the diesel fuel in the tank would therefore exceed the maximum permissible temperature. The excess fuel must therefore be cooled before being introduced into the tank.

The fuel can be cooled by means of an air-cooling or coolant-cooling system. Coolers operating along similar lines to oil-to-air coolers or disk-stack oil coolers are used to cool fuel.

Figure 12: Cooling module

1 Transmission-oil-to-coolant cooler,
2 Coolant radiator,
3 Intercooler,
4 Condenser,
5 Air-to-power-steering-fluid cooler,
6 Module frame,
7 Module bearing,
8 Electric fan,
9 Double fan cowl,
10 Transmission-oil lines.



Modularization

Cooling module

A cooling module is a structural unit which can consist of various components for cooling (e.g. main radiator, intercooler, oil-to-air cooler) and a condenser for air-conditioning a passenger car, and can include a fan unit with drive (e.g. electric motor) (Figure 12).

The design of a cooling module must take into consideration the interaction of the individual components, the dimensioning of the components with respect to the package space in the vehicle, and the procedure for dealing with interfaces. Important factors here are the mounting technology, the cooling-air ducts and the seals on the cooling-air side, the component connections on the fluid side, and the electrical plug-in connections.

The standard modular construction used in both passenger cars and commercial vehicles essentially offers a whole range of technical and economic benefits:

- Simplified logistics by combining components to form one structural unit,
- Reduced number of interfaces,
- Simplified mounting and assembly,
- Optimum component design by using matched components,
- Modular systems, encompassing various engine and equipment variants,
- Improved quality with regard to assembly.

Simulation and test methods are employed for the purpose of achieving optimized component design and layout in the cooling module. Assuming exact knowledge of the fan characteristics, the fan drive and the heat exchangers, simulation programs are created to replicate both the cooling-air side and the fluid side. By integrating the individual components into the simulation models, it is possible to examine the interactions of the individual components under various operating conditions. Ever greater significance is being attached to this type of virtual analysis, which is characterized by the use of computer-aided development tools. In line with this development, all geometric data are entered and processed in a CAD system (Computer-Aided Design). CFD analyses (Computational Fluid Dynamics) are con-

ducted to examine the flow of cooling air in the engine compartment, while FEM analyses (Finite-Element Method) provide statements concerning the strength and stability of the design layout. The design analysis phase concludes with verification tests that may also be performed in a wind tunnel and on vibration test rigs.

Cooling-system technology

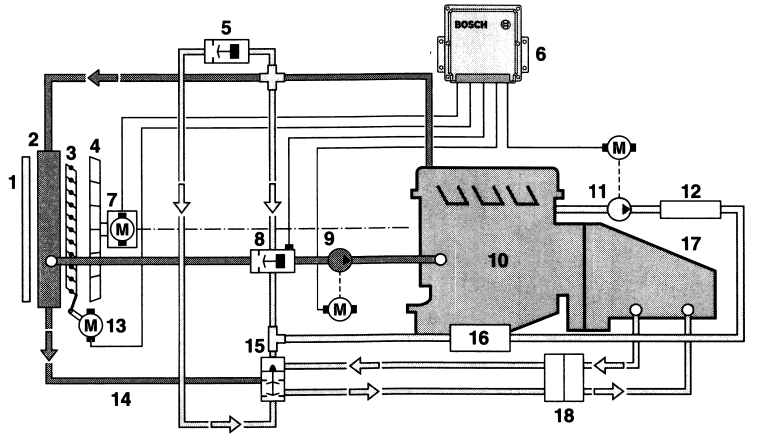
While the cooling module comprises a structural unit of components with defined functions, the cooling system encompasses all components that are associated with the functions of the cooling system, even if they do not form complete structural units (Figure 13). This includes for example, beyond the cooling module, lines, pumps, control components, and the expansion tank, if it is not already a component part of the cooling module.

Cooling-system technology, in which all the components are matched to each other, offers a whole range of technical and economic benefits:

- Reduced parasitic losses through appropriate hydraulic design matching,
- Consideration of control systems and dynamics,
- Consideration of the passenger-compartment heating system,
- Larger scope of intervention options for the purpose of optimizing the design,
- Standardized assembly concept for all cooling-system components,
- Reduced development expenditure by cutting the number of development interfaces.

Figure 13: Controlled cooling system

Example architecture with heat exchangers and actuators.
1 Condenser, 2 Coolant radiator, 3 Shutter, 4 Fan,
5 Thermostat, 6 ECU, 7 Electric motor (fan drive),
8 Electronically controlled thermostat,
9 Electric motor with pump, 10 Internal-combustion engine,
11 Electric motor with pump, 12 Heater core,
13 Stepper motor, 14 Low-temperature circuit,
15 Low-temperature regulator, 16 Oil cooler,
17 Transmission, 18 Transmission-oil cooler.



Intelligent thermal management

Function

Future trends are heading toward the system-optimized regulation of various heat and substance flows. Thermal management goes beyond cooling-system technology in that it takes into consideration all material and heat flow systems in the vehicle, i.e. in addition to the flow structures of the cooling system, it also deals with those of the air-conditioning system. The objectives of optimization consist in reducing fuel consumption and exhaust emissions, increasing air-conditioning comfort, increasing the service life of components, and improving cooling capacity in part-load states.

Optimization objectives

One of the basic principles of thermal management concerns itself with the fact that auxiliary energy employed to operate the cooling system always represents a loss for the vehicle energy balance, and component efficiency cannot be increased arbitrarily at a constant supply of auxiliary energy. To achieve the optimization objectives, the cooling system is therefore equipped with "intelligence", installed in familiar and new types of actuators, as well as in microprocessor-controlled control systems that operate these actuators. One example is demand-triggered regulation of cooling-air flow by radiator shutters and controllable fan drives so that cooling-air throughout is kept to a minimum under all operating conditions. In addition to improving the vehicle's drag coefficient c_d), this measure also ensures that all media achieve their operating temperature more effectively during the warm-up phase after a cold start, and that the passenger compartment is heated more efficiently. Cutting back the use of auxiliary energy in this way means that auxiliary energy can be diverted for use in operating states that are critical to the cooling output, while still achieving the optimization objectives.

Another important basic principle is to maintain a constant temperature in the components to be cooled as far as possible, irrespective of the operating state and the ambient conditions. An example of this temperature control principle is to use coolant to regulate the transmission oil temperature. Heating the transmission oil during the warm-up phase and employing an efficient cooling system to prevent the transmission oil from overheating reduce friction losses in the transmission, increase the service life of the transmission, and extend the service intervals for the transmission oil.

Ultimately, considering the cooling and air-conditioning systems in their entirety opens up the option of utilizing "thermal integration". Heat flow from one of the systems can be utilized or dissipated by another system without the need for any major additional input in auxiliary energy. An example of this is the utilization of waste heat from the exhaust-gas cooling system to heat up the vehicle interior.

Applications

In the field of engine cooling thermal management is used for the following:

- Transmission-oil temperature equalization,
- Map-controlled thermostat,
- Electrically controlled viscous clutch,
- Controllable electric coolant pump,
- Cooling-air control, e.g. radiator shutter,
- Exhaust-gas cooling,
- Coolant-cooled charge-air cooling (intercooling).

Advantages

The fuel-saving potential based on the sum of all measures is in the range of approx. 5% (for passenger vehicles). On top of this, there is a range of further advantages corresponding to the above-mentioned optimization objectives. The extent to which engine management utilizes the cooling system control options is of decisive significance in leveraging this potential.

In the meantime, individual measures have been implemented to achieve system-optimized temperature equalization in motor vehicles. Nevertheless, thermal management as an all-encompassing optimization principle has to date not been fully implemented and remains the reserve of future vehicle generations.

Engine lubrication

Force-feed lubrication system

Heat and wear are produced in engines by the friction of moving parts. Small amounts of residual dirt from production and particulate ingress from the outside are also to be expected in the engine system. The engine system is lubricated with lube oil in order to minimize friction and eliminate or reduce wear and to remove particulates and heat from the engine system.

The force-feed lubrication system in combination with splash and oil mist lubrication is the most commonly used system for lubricating motor-vehicle engines. An oil pump (usually a gear pump) conveys pressurized oil to all lubricating and bearing points in the engine (Figure 1), while the sliding parts are supplied with lube oil and lubricated by splash lubrication systems and oil mist.

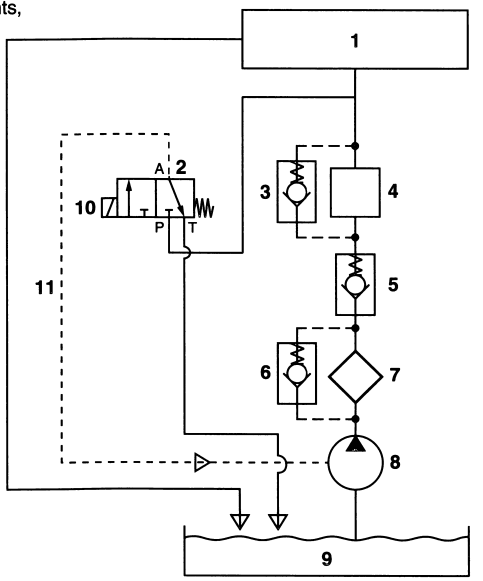
After flowing through the bearing surfaces and sliding parts, the oil collects below the engine in the oil pan. The oil pan is a reservoir where the oil cools, and the foam dissipates and settles. In the meantime an additional oil cooler is used in virtually all engines to regulate the oil temperature. In addition to filtration of the oil, thermal management is of particular importance.

Engine service life can be prolonged significantly by keeping the oil clean, i.e. filtration. Oil cooling by means of good thermal management makes possible current technologies such as single or multiple charge-air charging and the associated downsizing of engines.

Figure 1: Schematic diagram of a forced-feed lubrication system

(Schematic diagram).
1 Engine lubricating and bearing points,
2 Solenoid valve,
3 Cooler bypass valve,
4 Oil-to-water heat exchanger,
5 Anti-drainback valve,
6 Filter bypass valve,
7 Oil filter,
8 Variable oil pump,
9 Oil pan,
10 Signal from engine control unit,
11 Signal for pressure-stage preselection.

A Working line:
Pressure signal for oil pump.
P Pressure line:
Oil-gallery pressure.
T Ventilation line:
Ventilation to crankcase.



Components

Oil filters

Function

Oil filters remove and reduce particulates (soot and other residues from incomplete combustion, metal abrasion, dust, etc.) from the engine oil, which could otherwise cause damage or wear in the lubrication circuit. Wear in this context means that the particulates can get between the moving engine components, e.g. between piston and cylinder wall. Here the surfaces can be damaged by the particulates, which in turn can cause scoring or further particulate production. Since the engine oil is constantly circulating in the lubricating oil system, inadequate filtration could cause particulates to accumulate and this would accelerate the rate of wear of components that move against each other. The oil filter does not filter out liquid or soluble constituents such as water, additives or decay products attributed to the effects of oil aging.

The typical size of particles in engine oil ranges from 0.5 to 500 µm. The fineness of the oil filter is therefore specifically adapted to the requirements of a particular engine.

Different types and designs

In principle, oil filters are based on two specific designs: easy-change filters and housing filters. In an easy-change oil filter (Figure 2), the filter element is located in a housing that cannot be opened and which is secured by means of a threaded stud to the engine block. The complete easy-change filter unit is replaced as part of the oil service.

The housing oil filter (Figure 3) is made up of a housing that is permanently connected to the engine block and can be opened to access the replaceable filter element. During the oil service, only the filter element is replaced; the housing is a permanent component. The filter element as a rule is based on a metal-free design. This means that the filter element can be completely incinerated.

Figure 2: Easy-change oil filter

1 Threaded cap,
2 Filter element,
3 Filter bypass valve,
4 Spring,
5 Seal,
6 Anti-drainback diaphragm,
7 Housing
8 Center tube.

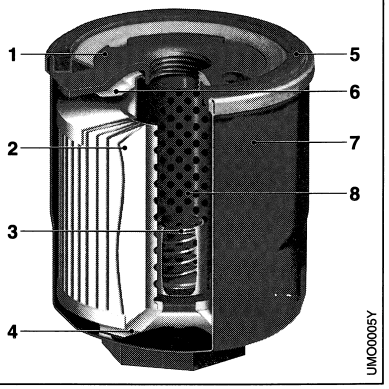
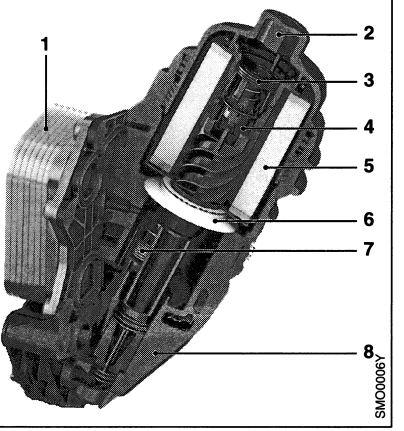


Figure 3: Housing oil filter

1 Oil-to-water heat exchanger,
2 Screw cap,
3 Filter bypass valve,
4 Center tube,
5 Filter element,
6 Anti-drainback diaphragm,
7 Anti-drainback valve,
8 Filter housing.



Besides the filter element, the two filter designs normally feature a filter bypass valve, which opens at high differential pressures to ensure effective lubrication at the necessary points in the engine. Typical opening pressures range from 0.8 to 2.5 bar. Elevated differential pressures can occur in connection with high oil viscosities or when the filter element is already heavily fouled.

Depending on the specific engine requirements, the two filter designs may also feature a non-return or backflow check valve on the filtered or unfiltered oil side (contaminated oil side). These valves prevent the oil filter housing from draining empty after the engine is turned off.

Currently, the oil and oil-filter change intervals for passenger-car engines are between 15,000 and 50,000 km, and for commercial vehicles are between 60,000 and 120,000 km – very much dependent on the engine system, place of use and the service strategy of the car manufacturer.

Filter media

Various types of deep-bed filter media are employed in oil filtration. These are entangled-fiber structures (nonwovens) which exist in different configurations. The filter material most commonly used is flat media which, in the majority of cases, is pleated, but in some applications they are also wound or used in the form of fiber packings, especially in bypass oil filters. A mixture of cellulose with synthetic fibers is the most widely used fiber material. The synthetic-fiber content can vary markedly depending on the requirements. Alternatively, amounts of glass fibers can be mixed with the cellulose. These filter media are impregnated with resin to enhance their resistance to oil and their operational stability.

However, filter media made of purely synthetic fibers are being used even more, as they have a much greater resistance to chemicals and clear benefits in terms of differential-pressure behavior. These attributes allow for longer service intervals. They also provide better options for structuring the three-dimensional fiber matrix to specifically optimize filtration and increase particulate retention efficiency.

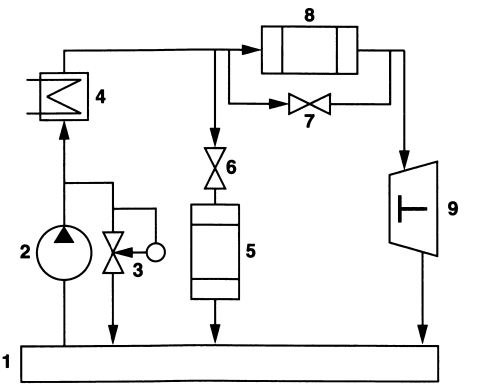
Full-flow oil filters

All vehicles are as a rule equipped with full-flow oil filters. Based on this filtration principle, the entire volumetric flow of oil that is pumped to the lubricating points in the engine is routed through the filter (Figure 4). Consequently, all particles which could cause damage and wear due to their size are trapped the first time they pass through the filter.

The decisive factors governing the filter area are oil volumetric flow and particulate retention capacity.

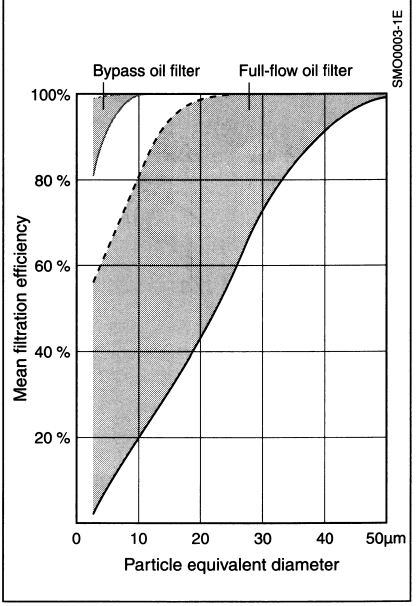
Figure 4: Oil circuit with full-flow and bypass oil filters

(Schematic diagram).
1 Oil pan,
2 Oil pump,
3 Pressure-control valve,
4 Oil cooler,
5 Bypass oil filter,
6 Throttle,
7 Bypass valve,
8 Full-flow oil filter,
9 Engine.



Bypass oil filters
Bypass oil filters, designed as deep-bed filters or as centrifuges, are used for superfine or microfiltration of the engine oil. These filters remove much finer particles from the oil than is possible using full-flow oil filters (Figure 5). They can remove minute abrasive particles to enhance wear protection. Soot particles are also filtered out to reduce any increase in oil viscosity. The maximum permissible soot concentration is roughly 3 to 5 %. Oil viscosity increases substantially at higher soot concentrations, resulting in a drop in the operational effectiveness of the oil. For this reason, bypass filters are used mainly on diesel engines. Only part of the oil flow (8 to 10 %) from the engine is routed via the bypass oil filter.

Figure 5: Filter fineness of full-flow and bypass oil filters (acc. to ISO 4548-12 [1])
------- Maximum filtration efficiency for bypass or full-flow oil filter.
——— Minimum filtration efficiency for bypass or full-flow oil filter.



Thermal management

Function

Aside from its lubricating function, the engine oil also has the task of transferring or carrying off heat. At the same time the engine-oil temperature must not exceed a critical value, as otherwise there is a risk of thermal destruction, so-called cracking of the oil.

Oil-to-water heat exchangers, which are frequently combined with the oil filter module, are used to control the temperature of the oil circuit. This combination allows for a more compact construction and the reduction of interfaces. The primary task is to transfer heat from the oil circuit to the water circuit. The heat energy is then dissipated via this to atmosphere for example with an air-to-water heat exchanger.

The heat transfer must be controlled to provide optimum thermal management. To quickly heat the engine oil after cold starting, the oil-to-water heat exchanger should only be activated after an increase in temperature in the oil. This can be done by thermostatic valves which are controlled with wax elements or thermo-springs. Electric solenoid valves provide for even better control options.

Oil-to-water heat exchanger

The preferred design for an oil-to-water heat exchanger is with plates stacked on top of one another. Oil and water flow alternately through the chambers. The heat is transferred via the plates to the neighboring fluid. To achieve high heat-transfer rates the flow in the columns should be as turbulent as possible. Structures are introduced into the heat-transfer plates or turbulence plates installed in the column for this purpose.

References
[1] ISO 4548-12: Methods of test for full-flow lubricating oil filters for internal combustion engines – Part 12: Filtration efficiency using particle counting and contaminant retention capacity.

Air-intake systems and intake manifolds

Overview

Functions

The air-intake system and the intake manifold are designed to supply the engine with air that is as cold and particle-free as possible, and to distribute this air to the individual cylinders. The main functions of the air-intake system are to direct the air from the front of the vehicle to the engine, to separate particulates contained in the intake air from the surrounding area, and to muffle the noises emitted by the engine.

Components

Basic functions

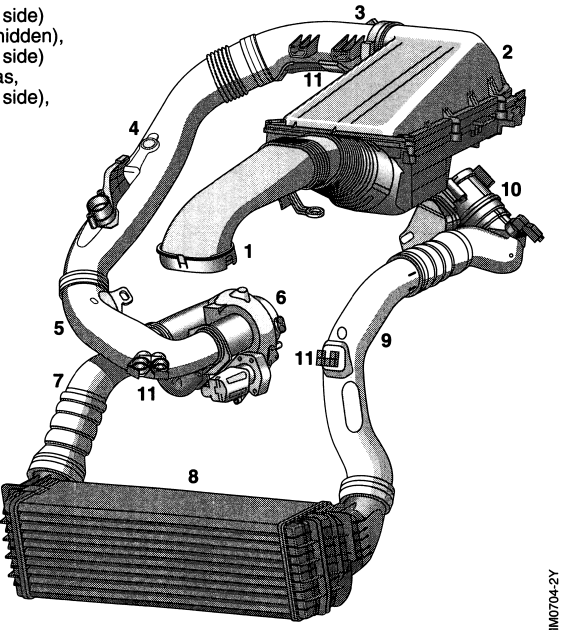
The air-intake system (Figure 1) consists of an air duct on the raw-air side, the air filter as the central components, and the lines to which the throttle valve and the intake manifold are connected. The intake manifold distributes the intake air to the individual cylinders and ensures uniform and effective combustion. The air filter prevents mineral dust and particles from being drawn into the engine and contaminating the engine oil. This reduces wear in the bearings, piston rings, cylinder walls, etc. In this way the individual system components can have a crucial positive effect on the performance and service life of the engine.

Supplementary functions

The air-intake system often incorporates further function components such as, for example, sensors (air-mass meter, pressure sensor, and temperature sensor), HC adsorber, water separator, and even anti-snow systems.

Figure 1: Design of an air-intake system with exhaust-gas turbocharging

1 Raw-air line,
2 Air-filter housing,
3 Clean-air line 1 (low-pressure side) with hot-film air-mass meter (hidden),
4 Clean-air line 2 (low-pressure side) with introduction of blow-by gas,
5 Clean-air line 3 (low-pressure side),
6 Exhaust-gas turbocharger,
7 Charge-air line (hot high-pressure side),
8 Intercooler,
9 Charge-air line (cold high-pressure side),
10 Throttle valve,
11 Attachment points.



Requirements

Functional requirements

The development of the air-intake system and its components is based on numerous specifications which are shaped on the one hand by space considerations and on the other hand by functional requirements. Specifications define for example temperatures in the engine compartment which influence the choice of plastic materials and elastomers for the individual components, permissible leakage rates and pressure losses for the system or the individual components, the acoustic performance of the intake tract, and the separation performance with regard to filtration as service life and as filtration-efficiency requirement.

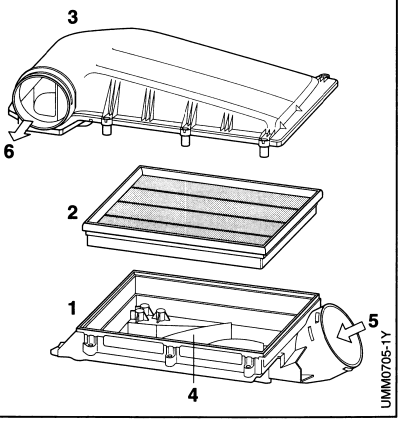
With regard to the intake manifold, further emphasis is placed on pressure-pulsation resistance and the level of uniform distribution of intake air. The air flow into the combustion chamber can if necessary be positively influenced by different types of flap.

Visual appearance

The visual appearance and the design of surfaces in the engine compartment gain in importance as soon as surfaces are located in the visible area of the engine compartment (e.g. design air filters or engine-mounted air filters).

Figure 2: Design of an air filter

1 Raw-air housing, 2 Air-filter element, 3 Clean-air hood, 4 Flow deflectors, 5 Raw-air connection, 6 Clean-air connection.



Passenger-car air-intake system

Air duct

The intake point of the raw-air line is usually located at the front of the vehicle. It is particularly important in that it has a crucial influence on the ingress of particles and the unwanted ingress of water droplets when it is raining and of snow crystals when it is snowing. However, this position is generally stipulated by the automobile manufacturers.

Classic internal-combustion engines are divided into non-supercharged naturally aspirated engines and supercharged engines, in which an exhaust-gas turbocharger for example is deployed. The air-intake system (Figure 1) in both variants consists of a raw-air line, which feeds into an air-filter housing with integrated filter element, and a clean-air line on the low-pressure side, which in a naturally aspirated engine serves as the interface to the throttle valve. In a turbocharged engine the clean-air line on the low-pressure side feeds into the turbocharger, while the charge-air lines on the hot and cold sides lead to the intercooler and to the throttle valve respectively. Low pressure losses can be achieved by large line cross-sections (this influences space and acoustics) and by optimized flow routing.

The lines are usually made of plastic and elastomer materials, while the air-filter housings are generally injected-molded plastic parts. Depending on the application, the air filter can be mounted to the body and to the engine in such a way that the lines compensate not only the tolerances between the components and their attachment points, but also above all the movement of the engine relative to the body. The components are joined with vibration isolation to the vehicle by way of rubber buffers.

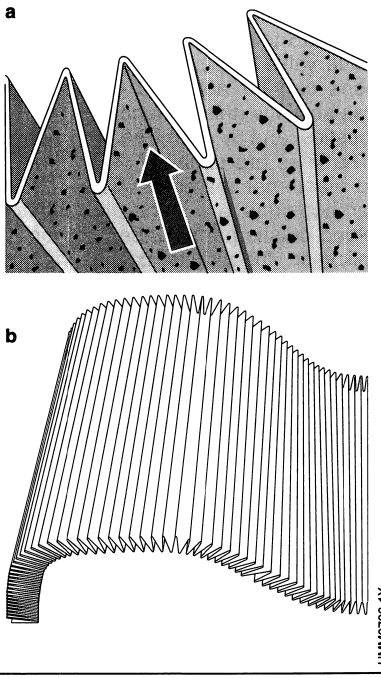
Air filtration

Engine air-intake filters (Figure 2) reduce the particles contained in the intake air which can be generated for example by abrasion and incomplete combustion or condensation processes or are even of natural origin in the form of organic and mineral dust.

Composition of air pollutants
Typical air pollutants includes oil mist, aerosols, diesel soot, industrial waste gases, pollen, and dust. These particles vary greatly in size. The dust particles drawn in together with the intake air have a diameter of between 0.01 µm (mostly soot particles) and 2 mm (sand grains). Approximately 75 % of the particles (referred to the mass (weight)) are between 5 µm and 100 µm.

The amount of dust contained in the intake air depends heavily on the environment in which the vehicle is used (e.g. freeway/interstate or dirt track). In extreme cases, over a period of ten years, the mass concentration accumulated in a passenger car can range between a few grams through to several kilograms of dust.

Figure 3: Filter media
a) Flow through a folded/pleated filter medium,
b) Folded/pleated filter medium.



Filter elements
Filter elements which incorporate the latest technology achieve total mass filtration rates of up to 99.8 %. It is necessary to maintain these values under all prevailing conditions – also under dynamic conditions as found in the engine's intake tract. Poor-quality filters will exhibit an increased rate of dust breakdown or rupture.

Design
Filter elements are designed to meet the requirements of each individual type of engine. This ensures that pressure losses remain minimal and that high levels of filtration efficiency are not dependent on air throughput. The filter medium that makes up the filter elements in flat or cylindrical filters is installed in folded or pleated layers in order to achieve a maximum filter surface area in the smallest possible space (Figure 3). Conical, oval, stepped, as well as trapezoidal geometries complement the standard structures with the aim of optimizing the use of package space that is becoming very scarce and confined in the engine compartment. These media, mostly consisting of cellulose fibers, go through a special embossing and impregnation process to achieve the necessary mechanical and thermal strength, sufficient water stability, and resistance to chemicals. Newer applications are equipped with flame-inhibiting filter media to reduce the risk of fire caused for example by an drawn-in cigarette.

The elements are designed in accordance with the spacings and replacement intervals defined by the vehicle manufacturer. The intervals are between two and four, and sometimes even six years, or 30,000 km up to 100,000 km.

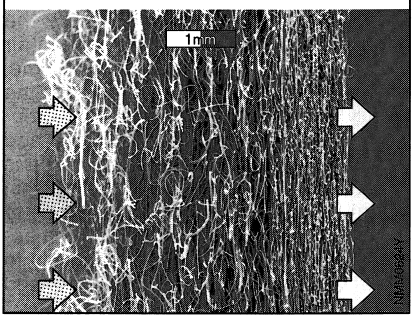
Filter media and design
The media used in passenger-car air filters are mostly part of deep-bed filters which, in contrast to surface filters, retain the particles in the structure of the filter medium. Deep-bed filters with a high dust absorption capacity are always used to advantage wherever an economic solution is required to filter large volumetric flow rates with low particulate concentrations.

The demands for compact, high-performance filter elements (with reduced package space), as well as longer service intervals, are the driving force behind the development of new, innovative air filter media. To further increase the dust retention capacity of deep-bed filters, design engineers are increasingly using media which have a gradient structure with increasing fiber density to the clean-air side.

Fully synthetic filter media
New air-filter media consisting of synthetic fibers with much improved performance data have already been phased into series production. Figure 4 shows a photograph of a synthetic high-performance filter medium (felt) with continuously increasing density and decreasing fiber diameter across the filter section from the input side to the output side.

Semi-synthetic filter media
Better results than with purely cellulose-based media can be achieved using composite materials. Here, for example, a cellulose-based paper layer is combined with a synthetic layer of melt-blown fibers. The melt-blown fibers constitute a layer of fibers formed from a polymer melt in the air flow which is either deposited directly on the paper layer or laminated on in a separate process step as a separate layer.

Figure 4: Photographic image of an air-filter medium consisting of synthetic fibers, taken with a scanning electron microscope

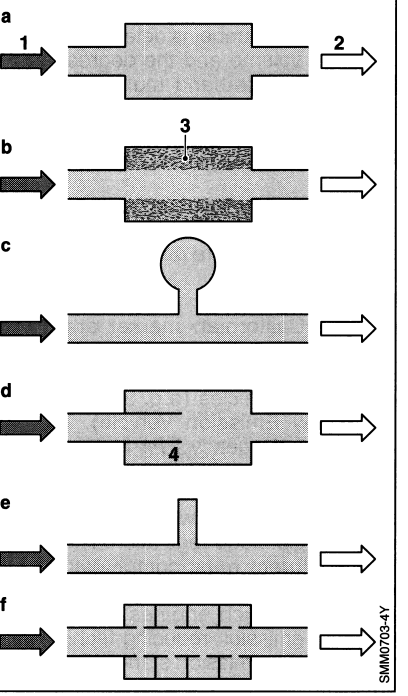


Acoustics

The air-intake system contributes in a similar way to the exhaust system to the overall noise of the vehicle. It is therefore necessary to introduce noise-reduction measures already in the early development stages. In the past these measures

Figure 5: Acoustic muffling measures with different resonator shapes

a) Resonator-type muffler (relatively wide-band, suitable for low and medium frequencies),
b) Absorption-type muffler (wide-band, suitable for medium and higher frequencies),
c) Resonance muffler (narrow-band, suitable for low and medium frequencies),
d) Whistle muffler (narrow-band, high muffling, suitable for medium frequencies above the Helmholtz resonance),
e) Branch whistle muffler (λ/4-pipe, narrow-band, suitable for medium frequencies),
f) Wide-band muffler for turbocharged engines (wide-band, suitable for higher frequencies).
1 Air duct, upstream,
2 Air duct, downstream,
3 Acoustically absorbent material,
4 Inserted pipe.



were confined to designing a sufficiently sized air-filter housing as an intake muffler. In the meantime the excitation spectrum has shifted from low frequencies (f < 200 Hz) to medium to high frequencies with the result that separate measures in the form of Helmholtz resonators and $\lambda/4$ pipes are used. Computer simulation programs that determine the optimum number of resonators and their frequency are used to design these measures.

Wide-band mufflers
for turbocharged engines
The increasingly widespread use of turbocharged engines has led design engineers to focus their attention on new noise sources. The whistling and whining noises are subjectively perceived to be disruptive and annoying and now measures are often integrated in the form of wide-band mufflers in the clean-air line of the air-filter system. Wide-band mufflers comprise several individual chambers through which a perforated central pipe passes (Figure 5). The individual resonant frequency of a chamber is determined by the chamber volume and the degree of perforation; a wide-band muffling effect is achieved by arranging several chambers in succession. Here too modern simulation programs are used to optimize the number.

In addition to resonator-type mufflers, absorption-type mufflers as featured in exhaust systems are occasionally used.

HC adsorption
For the Californian market and some other US states limits regarding the emission of highly volatile hydrocarbons are defined for vehicles (e.g. SULEV, Super Ultra Low Emission Vehicle). The US environment agency EPA is responsible for implementing the legal requirements regarding HC emission limits (High Volatile Hydro Carbons) within the framework of a multiple-stage regulation. The engine manufacturers must comply with these statutory provisions and adhere to time limits after which engines must feature pollutant-emission-reducing technologies to satisfy the emission standards.

Some of the unburned hydrocarbons (primarily fuel components) escape after the engine is turned off from the engine through the intake manifold and from the crankcase ventilation into the air-intake system, travel against the normal direction of air flow, and exit the system to atmosphere. This can mean that it is necessary to integrate an HC adsorber on the clean-air side of the system so that the hydrocarbons are separated by adsorption on activated-carbon media. An equivalent, the "Butane Working Capacity" (BWC), is defined in specifications in order to define the absorption capacity comparably. The adsorber elements can be used in the "full flow" (main path, full throughflow) or in a bypass configuration.

Unlike air-filter elements, adsorber elements according to the requirements are lifetime components which may not be removed from the system. The elements are charged while the engine is stopped with the air-borne hydrocarbons of minimal droplet size, and desorption takes place the next time the engine is operated, i.e. the hydrocarbons are returned to the combustion process in this cycle and the HC adsorber is regenerated for the next charging cycle. The adsorptive medium used is activated carbon, which is located as a single or double layer between the carrier media. These are for the most part elements which are manufactured in a process as plastic coating.

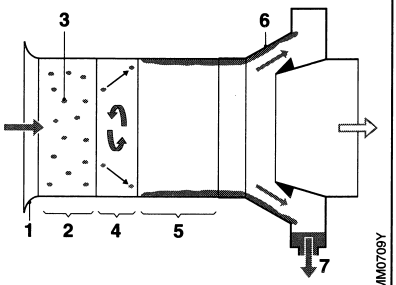
Water separation
When driving in the rain, not inconsiderable amounts of water may be drawn in, depending on the position and layout of the intake fitting at the front of the vehicle. Because the filter medium only has a limited water absorption capacity, the drawn-in liquid may reach the clean-air side of the system after a certain time. Signal deviations or even damage to the hot-film air-mass meter may occur if droplets touch the hot sensor surface and result in a local, unwanted cooling of this surface. The signal deviation suggests an incorrect air-mass flow and gives rise to an incorrect setting of the mixture composition, to power losses, and ultimately to increased consumption (and increased

CO_2 emissions). The time until water penetration at the filter element can be significantly extended or even avoided entirely by way of upstream measures.

Finned separators, baffle plates or measures based on the cyclone principle such as the "peeling collar" (Figure 6) integrated in the intake air are used to separate water droplets. The shorter the raw-air line – i.e. the distance from the air inlet to the filter element – the more difficult the task. Essentially, the aim is to achieve additional flow pressure losses that are as low as possible. The measures must be combined with suitable water-discharge devices which divert the separated water from the system to atmosphere. However, separating principles are also used which carry the liquid in combination with measures on the housing side and with additional water-separating fleeces attached to the filter element from the system.

An optimized design of the raw-air line and the air-filter housing clearly enhance the above-mentioned measures such that outstanding separation results are achieved even with comparatively short raw-air lines.

Figure 6: Principle of water separation
1 Raw-air intake,
2 Raw-air line,
3 Water droplets,
4 Swirl generator (optional),
5 Wall-applied water film,
6 Peeling collar,
7 Water discharge.

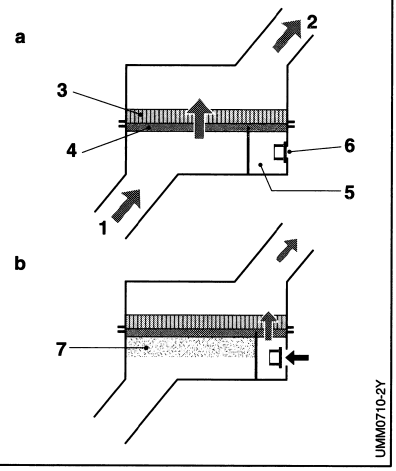


Anti-snow system

In winter when it is snowing snow crystals may be drawn into the system which charge the filter element on the raw-air side. If this situation persists, the initial upshot will be an increase in pressure loss. In due course the filter element may become blocked on the raw-air side such that the engine stops receiving sufficient intake air for the combustion process. This will ultimately cause the engine to stall.

To prevent this, anti-snow systems (ASS) – essentially a second intake – are used (Figure 7). If the filter element becomes blocked, hot, dry and snow-free air is drawn in from the engine compartment, whereby this reduced air flow maintains the engine function. The supplementary intake can be opened by temperature- or pressure-controlled control mechanisms. In the case of thermal control a flap is moved by an expanding-wax element. Pressure control is based on a spring-loaded valve integrated in the

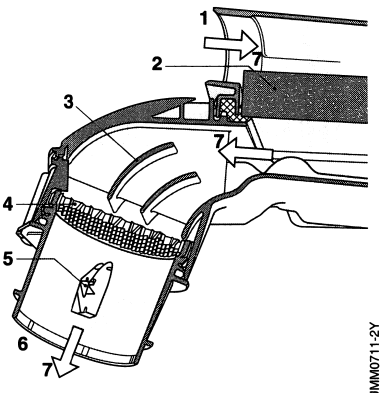
Figure 7: Anti-snow system
a) Bypass valve closed,
b) Bypass valve open.
1 Raw air,
2 Clean air,
3 Air-filter element,
4 Fleece on raw-air side,
5 Dry chamber,
6 Bypass valve,
7 Snow block.



air filter on the raw-air side. The filter element is basically fitted with a preliminary fleece which assumes the sealing function and protects the "dry" chamber of the second intake against snow ingress.

Sensor technology
Sensors which measure the pressure, the temperature and even the air-mass flow are often integrated on the clean-air side of the air filter. The exhaust-gas-relevant air-mass flow is measured by a hot-film air-mass meter. As the emission-control legislation is tightened the requirements with regarding to signal quality are becoming increasingly stringent. Signal deviations for an air-filter system in the range of ±2.5% are difficult to achieve, but not unusual. This requires an increased simulation outlay to identify disruptive turbulent areas and instabilities in the flow close to the sensor and to minimize them by means of suitable product design and additional measures or components (e.g. air-duct grills, fins or air deflectors) (Figure 8).

Figure 8: Flow optimization upstream of the hot-film air-mass meter
1 Raw-air connection,
2 Air-filter element,
3 Deflector geometry,
4 Rectifying grill,
5 Hot-film air-mass meter,
6 Clean-air connection,
7 Flow direction.



Visual design of surfaces
The structuring of plastic surfaces in the visible area called for by customers for visual reasons can be achieved in a variety of ways, such as for example by erosion, blasting or photoetching of the injection-molding die. Engine-mounted design air filters are a special example here.

Interfaces
The type of connection point between the air-filter housing and the clean-air line is often geared towards customer requirements. The production plant of the automobile manufacturer often specifies the shape of the interface. If an elastomer component is called for, lines based on the material PP-EPDM with integrated bellows for isolation are manufactured by injection molding. These can be connected to a line section made using blown-part technology.

Crankcase ventilation
The clean-air line on the low-pressure side can incorporate further interfaces such as for example the blow-by introduction point from the crankcase ventilation or even the extraction point for the PCV (Positive Crankcase Ventilation). The blow-by introduction point is frequently fitted with a heating tube which is intended to prevent icing.

Low-pressure exhaust-gas recirculation
A measure specially suited to smaller diesel engines to meet the strict Euro 6 limits for nitrogen oxides (NO_x) is low-pressure exhaust-gas recirculation (see Exhaust-gas recirculation, EGR). But even larger-sized engines utilize low-pressure exhaust-gas recirculation, often in conjunction with SCR systems (Selective Catalytic Reduction). With low-pressure exhaust-gas recirculation the exhaust gas is diverted after the diesel particulate filter, cooled in an EGR cooler, and introduced into the clean-air line upstream of the turbocharger compressor. To obtain high EGR rates in wide ranges of the engine map, it is very important for the EGR system to have a low-pressure-loss

design including mixing of air with exhaust gas. For special operating points in the part-load range an additional pressure drop is generated in the EGR system by an exhaust-gas restrictor (downstream of the EGR extraction point) or by an air-intake restrictor in the clean-air line (upstream of the EGR introduction point). Configuring the mixer geometry appropriately prevents local temperature gradients from damaging the compressor and compressor efficiency from being reduced. After the EGR cooler the exhaust gas reaches maximum temperatures of 150 to 200 °C, which enable the mixing and control functions to be executed cost-effectively and space-efficiently with the aid of plastic components. However, it is important when choosing the material to ensure that it is resistant to acidic condensate (pH value of 2 to 3) which can be produced at exhaust-gas temperatures below 85 °C.

The use of low-pressure exhaust-gas recirculation is being analyzed even in gasoline engines; however, the objectives here are to reduce CO_2 and the knock tendency rather than to reduce NO_x emissions. The first applications in series production are expected in the next few years.

Validation of air filter

Full product validation, consisting of simulation, acoustics and component testing, is an essential part of the development process for air-filter systems in order to achieve a high degree of product maturity as early as possible. Already in the early concept phase – within the framework of virtual product validation – simulations are conducted to optimize the functions and properties of air-filter systems. Here the acoustic properties such as intake-opening noise and surface sound radiation are calculated and optimized.

As well as acoustic behavior the fluid properties such as pressure loss and approaching flow to the measuring-sensor system are analyzed and specifically improved. In addition, the vibration characteristics are tested under roadway or engine excitation using finite-element simulations to fulfill the later life test without any problems. The filtration properties among others are checked as soon as the first sample components are available. The scope of testing within the framework of component testing includes tests on individual components and on the complete system. This is rounded off by functional tests where the ambient conditions prevailing in the vehicle (such as temperature, moisture, and media pressurization) are superimposed.

The component life is tested in comprehensive vibration and pulsation tests which simulate the real load in the vehicle. In view of the strict emission-control legislation, the flow behavior in the air filter must also be very closely analyzed with a dust charge also taken into consideration. Influencing factors such as fold/pleat geometry of the air-filter element and spread of the filter medium are taken into consideration here. Component validation is rounded off by simulations of the injection-molding process to optimize component quality and the manufacturing process.

Trends – further development of air-intake systems

Three main trends can currently be observed in the market: Downsizing of engines, platform strategies to utilize synergy effects, and stricter requirements with regard to emission-control legislation.

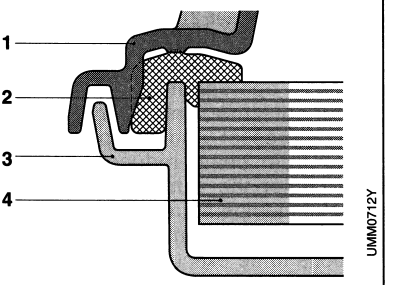
Downsizing necessitates a higher proportion of supercharged engines with the following effects: Specifically higher air-mass flows, smaller and more complex air-filter housings, increased conflict between acoustics and pressure loss, higher temperatures in the engine compartment, and higher charge-air pressures. This necessitates a higher simulation outlay for example to optimize flow and acoustics, new air-filter media, innovative filter-element designs, and higher-quality plastic materials.

Platform strategies of the automobile manufacturers necessitate more complex package spaces and a high-grade identical-part strategy. Stricter requirements

with regard to emission-control legislation give rise to heightened specifications with regard to sensor technology, particularly the quality of the signal stability of the hot-film air-mass meter. This likewise necessitates more complex simulation tools, new test procedures, and more complex product designs and constructions.

The weights of components can be reduced with new production technologies. Air filters produced by injection molding on the basis of foamed plastics help to achieve this. A further possibility is to reduce wall thicknesses, where the acoustic behavior in particular of the air-filter housing must be observed. Structure-borne noise radiation can be counter-compensated by cambering (curved geometry) surfaces and integrating ribs or beads (groove-shaped depressions), even when smaller wall thicknesses are used. New sealing concepts for the interface between filter element and housing, e.g. designed as a radial concept (Figure 9), render metal compounds superfluous, and result on the one hand in reduced costs and on the other hand in increased service friendliness when replacing the filter element. The two housing sections can be connected to each other by injected snap-on hooks.

Figure 9: Radial sealing principle
1 Air-filter bellows,
2 Radially acting seal,
3 Raw-air housing,
4 Clean-air hood.



Hybrid vehicles and active noise measures

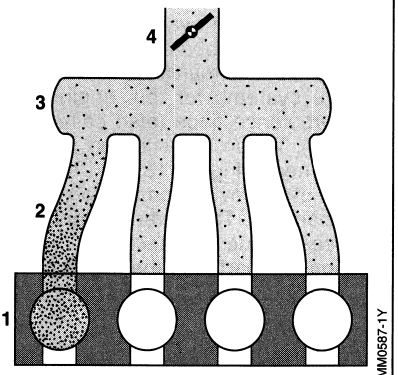
In the years to come hybridization in the automotive industry will present new challenges. Switches in operation between the internal-combustion engine and the electric motor, but also purely electric operation generate unusual acoustic states. Active systems (with ANC, Active Noise Control) constitute an interesting remedial measure in that the can not only reduce noise and thus replace passive components, but also create new degrees of freedom with regard to noise generation, e.g. for warning pedestrians in purely electric operation. Some countries have already initiated legal initiatives prescribing a warning function for purely electric operation. There are also opportunities to support a specific marque sound or for noise control in order to acoustically mask unwanted phenomena. This approach makes active systems interesting for future engine concepts.

Passenger-car intake manifolds

The main function of the intake manifold is to distribute the air homogeneously to the individual cylinders in order to ensure combustion that is as uniform and effective as possible in the individual cylinders. The quality of uniform distribution has a decisive effect on the engine's performance and emission characteristics.

In the case of gasoline engines a distinction is made between naturally aspirated engines and supercharged engines; in the case of diesel engines, on the other hand, only supercharged engines are still relevant. All three categories feature both passive and active intake manifolds. Active intake manifolds are usually called variable induction systems.

Figure 10: Ram-tube supercharging
1 Cylinder,
2 Individual ram tube,
3 Plenum chamber,
4 Throttle valve.



Intake manifolds for naturally aspirated engines

In these types of intake manifolds gas-dynamic supercharging effects are used to increase the air mass in the cylinder and thus the engine output or torque. There are two different supercharging effects here.

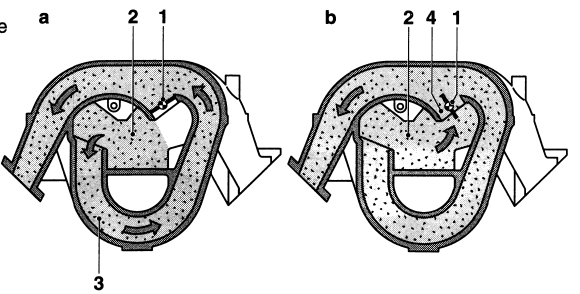
Ram-tube supercharging

Ram-tube supercharging is based on the principle that each cylinder has its own ram tube that feeds into a common plenum chamber, also called the intake plenum (Figure 10).

The intake pulse in the induction stroke triggers a vacuum wave. This propagates in the ram tube, is reflected at the opening to the intake plenum, and returns as a pressure wave to the combustion chamber. Because pressure waves always propagate at the speed of sound, the length of the ram tubes can be used to determine the engine-speed range in which the supercharging effect is wanted.

To be able to utilize this supercharging effect in numerous engine-speed ranges, there is the option of realizing active intake manifolds of this type (variable induction system, Figure 11). The optimum solution

Figure 11: Variable induction system
a) Intake-manifold geometry with changeover flap closed (torque setting),
b) Intake-manifold geometry with changeover flap open (power setting).
1 Changeover flap,
2 Plenum chamber,
3 Long, thin oscillatory intake passage with changeover flap closed,
4 Short, wide oscillatory intake passage with changeover flap open.



would be an infinitely variable induction system in which an optimum ram-tube length is assigned to each engine-speed range. But this is barely feasible, both technically and economically. For this reason mostly two- or in rare cases also three-stage induction systems are on the market.

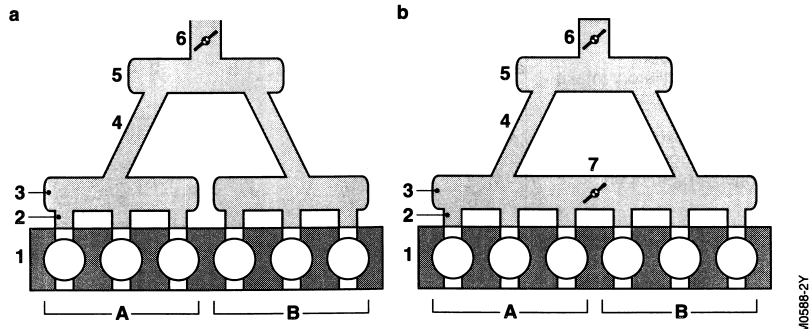
Tuned-intake-tube charging
Tuned-intake-tube charging is based on the principles of a Helmholtz resonator. Two groups of cylinders with the same ignition interval are established. The reason for this is that the intake pulses within the groups must not overlap. These groups are connected by short intake manifolds to the two resonance chambers (resonance plenums) (Figure 12). The two resonance plenums are connected by resonance (tuned) tubes to the common main intake plenum. The periodic intake pulses induce the gas oscillations to resonate at a specific engine speed and thereby generate the supercharging effect. The engine-speed range in which the supercharging effect acts is determined by the length and the diameter of the tuned intake tubes. This principle is preferably used in 3-, 6- and 12-cylinder engines because it is relatively easy, because of the firing order, to create independently operating cylinder groups (each consisting of three cylinders with 240° ignition interval).

Incorporating a resonance flap between the two resonance chambers creates an active system that provides further advantages. The intake manifold with closed resonance flap generates a torque increase in the lower engine-speed range. In the upper speed range this valve is opened to create an intake manifold with short ram tubes which increases power output in this speed range.

Intake manifolds for supercharged gasoline engines
Intake manifolds for supercharged gasoline engines are in most cases passive intake manifolds. Power and torque are primarily determined via the charge-air pressure generated by the charger and thus by the air mass introduced into the cylinders. The still available, but mostly very short ram tubes are typically used for passive intake manifolds.

Active systems for this engine type have turbulence flaps (also called tumble flaps). These are situated near the cylinder head (Figure 13) and have the function of improving combustion in part-load operation. This arrangement firstly increases torque and secondly reduces emissions. In part load the flap is closed

Figure 12: Tuned-intake-tube charging
a) Resonance induction system, b) Resonance variable induction system.
1 Cylinder, 2 Short intake manifold, 3 Resonance chamber,
4 Resonance tube, 5 Plenum chamber (main intake plenum),
6 Throttle valve, 7 Resonance flap.
A, B Cylinder groups.



and generates a cylinder flow about the horizontal axis (tumble). In the power setting the flap is opened and the engine receives through the opening cross-section as much air mass as possible into the cylinder.

A special form here is the trough flap, which in the open setting disappears completely into the duct wall and thus ensures unrestricted flow into the cylinder.

Intake manifolds for supercharged diesel engines

In the case of intake manifolds for supercharged diesel engines there is no such clear distinction between passive and active variants. These intake manifolds practically do not have manifolds any more. All that remain are short inlet tulips that connect the intake plenum to the cylinder head.

In the case of active intake manifolds the cylinder head always has two intake ducts per cylinder, where one is a swirl duct and the other a filling duct (Figure 14). The filling duct is fitted in the intake manifold with a duct shutoff flap. The function of this flap is to fill the cylinder via the swirl

Figure 13: Tumble arrangement
a) Intake manifold in tumble setting,
b) Intake manifold in power setting.
1 Intake manifold,
2 Turbulence flap (tumble flap).

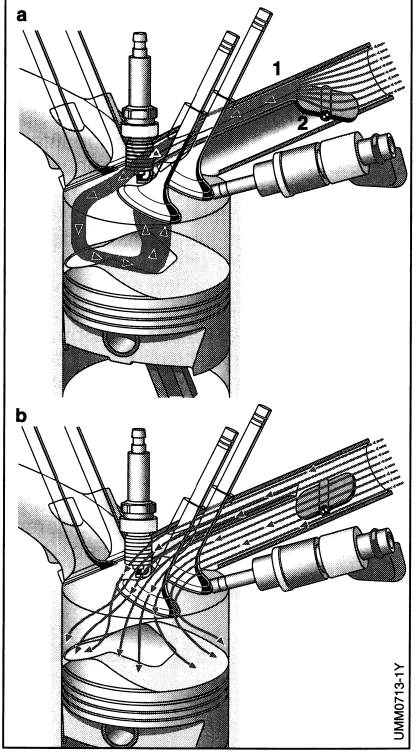
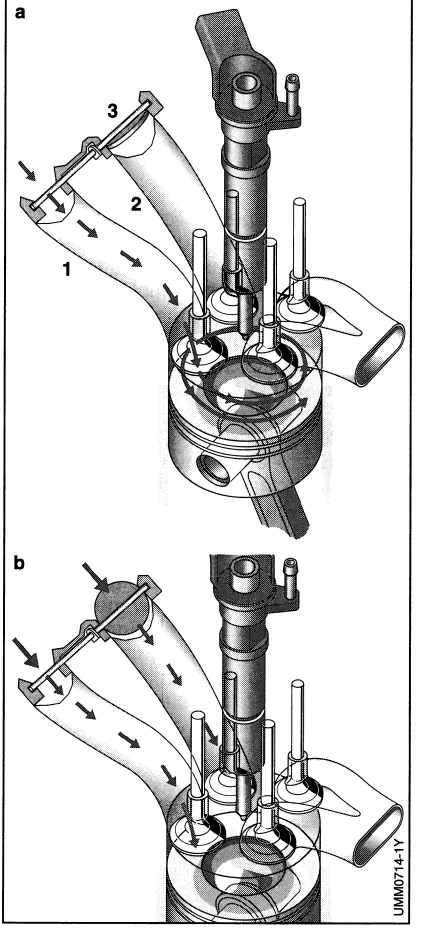


Figure 14: Swirl arrangement
a) Intake manifold in swirl setting,
b) Intake manifold in power setting.
1 Swirl duct,
2 Filling duct,
3 Duct shutoff flap.



duct in part-load operation and thus create a pronounced swirl about the vertical axis. This swirl ensures that air and fuel are thoroughly mixed and thereby helps to reduce emissions. In the transition to full-load operation the flap is then opened incrementally and the swirl is nullified by the filling duct. This results in as much air mass as possible being delivered to the cylinder and thus in high torque and high power output.

Actuator and sensor technology
Primarily electrical actuators or vacuum units are used to drive the different switching elements.

Electric actuators
Electric actuators are preferably used to activate swirl flaps in diesel engines. They enable any intermediate settings to be accurately and rapidly approached. Control is provided in most cases by the engine ECU.

Vacuum units (pneumatic actuators)
A vacuum unit can be used for pure two-point switchings, as are sufficient for length, resonance, and tumble switchings. These are switched by means of an electric changeover valve. The prerequisite is that the engine or a vacuum pump makes sufficient vacuum available.

Sensor technology
The sensor technology that features in variable induction systems has increased significantly in response to the OBD II legislation, which states that all exhaust-gas-relevant switching organs must be monitored by the engine ECU and that the driver must be alerted if they malfunction. Exhaust-gas-relevant switching elements are the turbulence flaps (tumble flaps in gasoline engines, duct shutoff flaps in diesel engines). Switching elements for length or resonance switching are not included.

For electric actuators the internal angle-of-rotation sensors of the actuators can be used in most cases. For vacuum units a separate sensor is required, but this can also be integrated in the vacuum unit.

For active intake manifolds, in addition to the sensor technology for angle-of-rotation monitoring, pressure, temperature, or combined pressure/temperature sensors are mounted on the intake manifolds. These are necessary to supply the engine ECU with the information from the intake manifold required for combustion.

Introduction of gases
A further function of the intake manifold is to distribute introduced gases uniformly into the respective cylinders. In gasoline engines these are primarily crankcase-ventilation and tank-ventilation gases. Fittings are attached to the intake manifold for these gases and then connected by the engine manufacturer to the relevant hoses. When crankcase blow-by gases are introduced, an additional heating tube may be required to avoid icing.

In diesel engines it is important for high-pressure exhaust-gas recirculation to be correctly introduced. Often uncooled exhaust gas is introduced into the intake manifold during the cold-start phase or diesel particulate-filter regeneration. Suitable shielding measures must be taken here where plastic intake manifolds are used. A central introduction point with adequate mixing section is beneficial to the extremely important uniform distribution here. The uniform-distribution requirement is very difficult to achieve in cylinder-selective introduction systems where the gas to be introduced (e.g. from exhaust-gas recirculation) is introduced via a distributor line with a separate introduction point for each cylinder (i.e. each cylinder is supplied with gas via a separate introduction point) at the very different operating points in the engine map.

Gaskets
A trend towards higher-quality materials can be discerned in gaskets. This is mainly because low-pressure exhaust-gas recirculation is more frequently used in diesel engines and this results in an aggressive condensate in the induction system. Peroxide-cured fluorocarbon rubber, for example, must be used in such applications. Due to the fact that the operating temperatures can extend both up and down, sometimes special low-temperature materials are required.

Validation of intake manifolds
Virtual and experimental product validation plays an increasingly important role for the intake manifold. The foundations for a flow- and weight-optimized component are already laid in the early development phase by the close interplay between design, testing and simulation. The optimum uniform distribution of air to the individual cylinders is optimized by CFD simulations in a number of iterative loops. The intake-manifold geometry is optimized with the aid of FEM simulations in such a way that maximum rigidity and strength are achieved with minimum weight. This enables the use of materials for the component to be kept as minimal as possible. The manufacturability of the component is analyzed and optimized with the aid of injection-molding simulations.
These theoretical analyses are then verified on sample and standard parts. The uniform distribution of intake air to the individual cylinders is measured on a flow-rate test rig. The structural strength of the components is ensured by rigorous pressure-pulsation and vibration tests under temperature influence. In the case of variable induction systems the durability of the switch-over elements and actuators is substantiated with endurance and fatigue tests.

Development trends for intake manifolds
Due to the switch from direct to indirect intercooling and the resulting benefits with regard to CO_2 emissions, the main trend in the intake-manifold market is towards the integration of intercoolers.

This trend has above all an effect on the intake manifolds in supercharged gasoline engines. Virtually all manufacturers pursue solutions involving full integration of the intercooler in the induction system. The cuboid-shaped intercooler must be integrated in the plastic intake manifold and, despite the large flat surfaces and sometimes for plastic extremely high temperatures (upstream of the intercooler), satisfy the requirements with regard to pressure-pulsation resistance. Materials that meet these requirements are currently in development.

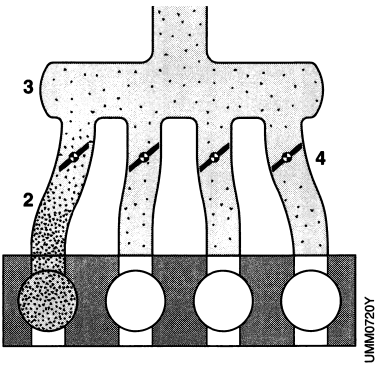
The diesel-engine market is split two ways: firstly the intercooler directly integrated in the intake manifold, and secondly the arrangement of the intercooler as a "single box" upstream of the throttle valve.

Motorcycle intake manifolds

Essentially, intake manifolds for motorcycle engines have the same functions as those for passenger-car engines. However, they do have the following particularities in terms of design: While in passenger-car applications one joint throttle valve is for the most part used for a number of cylinders (see Figure 10), for motorcycles frequently individual throttle valves are used (Figure 15). This permits a significant reduction of the intake-manifold volume between throttle valve and intake valves. The smaller volume increases the dynamics and thus the response in the event of both sudden throttle-valve opening and throttle-valve closing. The dynamic response is an elementary requirement in motorcycles. The geometry permits a smaller ratio of throttled intake-manifold volume to individual-cylinder displacement. This results in, aside from an improved response, intake-manifold pressure dynamics (see Cylinder charge).

The individual throttle valves can in this design be mounted on a single shaft or on a number of shafts (e.g., in V-engines). Comparable intake-manifold configurations are used in passenger-car engines, e.g., in racing engines.

Figure 15: Motorcycle intake manifold.
1 Cylinder, 2 Individual intake tract,
3 Plenum chamber, 4 Throttle valve.



Commercial-vehicle air-intake system

Air duct

The design of the air-intake system for commercial vehicles is essentially comparable to that of passenger-car air-intake systems (see Figure 1). What must be taken into account here is the much higher service life of commercial vehicles, the higher rated volumetric flow, and the broad range of uses in local, long-distance and construction-site applications. The different volumes and types of dust must be taken into account in filter design.

Raw-air duct, air filters and clean-air pipes up to the turbocharger are made of thermoplastics. After the turbocharger for the most part come (still) combinations of aluminum and rubber parts. However, more plastic parts will be used in future for charge-air pipes and intake manifolds (boost manifolds).

Raw-air intake

The positioning of the intake point is extremely important. A roof or side intake is usually chosen on heavy commercial vehicles. Both are generally mounted with rubber mounts for acoustic isolation to the cab rear wall. The dust and water concentrations in the air are lowest with the roof intake. There are however many raw-air duct variants to cater for the different driver's-cab heights and widths. A good compromise in terms of costs is the side intake (Figure 16): Here air is still inducted from height at the side rear of the driver's cab, irrespective of the different roof heights.

A low-height front intake or side intake, along similar lines to that used in passenger cars, is usually chosen for light utility vehicles. To keep the volumes of inducted dust a water in check, air is inducted through baffles on the vehicle sides. An air duct that delivers the best flow possible must be provided in all cases. Hot engine waste air must not be inducted as this mean a performance drop and higher fuel consumption.

Because most truck cabs are tilted forwards to access the engine for servicing, the raw-air duct is connected via bellows

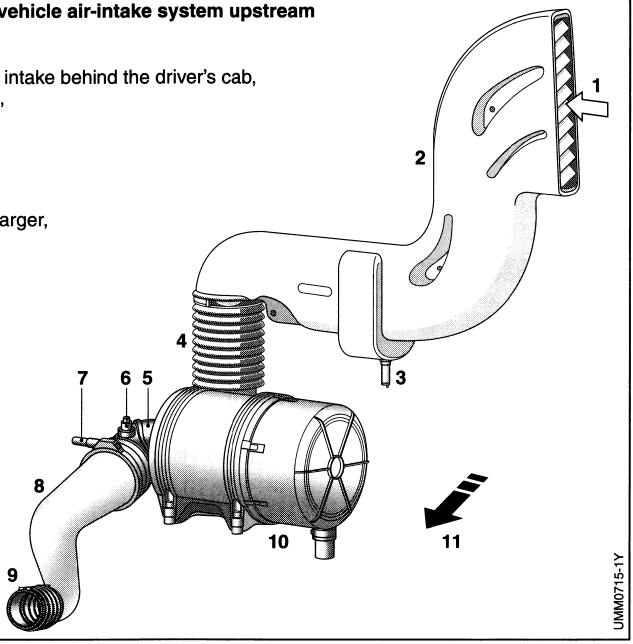
to the rest of the system. The function of the bellows is to seal the interface when the cab is tilted back and to compensate the vibrating motion of the sprung cab.

Water separation
Measures for water separation are integrated in the raw-air system. Water separation prevents dust from being washed through the filter medium, which can increase engine wear, but also cause sensors to malfunction. Because as a rule cellulose-based filter media are used, at worst the filter element may be destroyed. Measures for water separation include for example using a large intake cross-section to reduce the intake velocity so that basically as little water as possible is inducted. Baffles in the system cause the water droplets in the intake air to be thrown against the inner manifold walls, where they form a film. The water droplets are then separated via a peeling collar or a change in volume and discharged via a water-discharge valve.

Dust preseparation
In vehicle applications with a higher dust concentration in the air (e.g. construction sites or operation in dusty countries) a dust preseparator is installed in the raw-air supply, either as an individual cyclone (Figure 17) or as a cell cyclone. Cell cyclones are small cyclones which are integrated in a housing to save space. These separate out a high proportion of the coarse dust contained in the intake air. The air is induced to rotate at the inlet by guide vanes; the heavy dust particles are thrown against the inner wall of the housing and discharged through the dust outlet. The precleaned air is directed by a further guide-vane assembly (optional), which rectifies the air flow again, to the air filter.

Air filters
The most widely used air filters in commercial vehicles at present are single-stage round air filters with spiral vee-shaped filter elements. They are characterized by

Figure 16: Commercial-vehicle air-intake system upstream of the turbocharger
1 Inlet grill,
2 Raw-air duct with side intake behind the driver's cab,
3 Water-discharge valve,
4 Bellows,
5 Rubber elbow,
6 Maintenance switch,
7 Screw clamp,
8 Clean-air pipe,
9 Connection to turbocharger,
10 Air filter,
11 Direction of travel.



low pressure loss, high dust capacity, and thus long service intervals. Depending on how they are used, they can extend a typical oil-change interval by a factor of two. A further advantage of round filters is that they can also be fashioned as two-stage filters. In these variations dust separation takes place inside the air filter itself (Figure 18). The tangential air inlet generates a swirl in the housing which separates out the coarse particles.

Depending on the application, different filter media are available today which are optimized for example for maximum filtration efficiency of fine dust and soot.

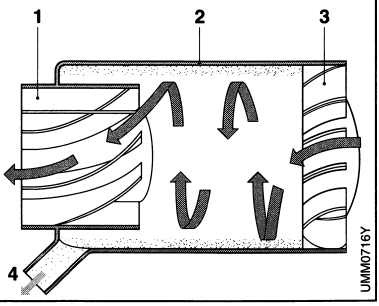
A safety element is optionally fitted inside the main filter element which prevents dust from reaching the clean-air side during servicing. It also protects the engine even if the main filter element has been damaged during servicing. Basically, however, air-filter elements should be replaced and not cleaned.

To optimize vehicle acoustics, the air filter is also fitted where necessary with resonators to minimize the opening or engine-braking noise.

Clean-air line as connection to the turbocharger

The clean-air line is as a rule a combination of metal, plastic, and rubber parts. Sensors for air mass, vacuum and air moisture can be optionally integrated. It is important for all the interfaces on the clean-air side to be permanently tight. The vacuum switch indicates to the driver that the servicing limit for the filter element has been reached. Flexible rubber connections compensate the relative movement between the engine and the frame-mounted air filter.

Figure 17: Dust preseparator (cyclone)
1 Outlet guide vanes, 2 Housing,
3 Inlet guide vanes, 4 Dust outlet.



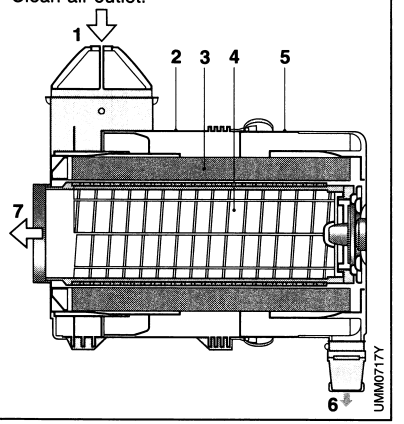
Air filtration

Filter elements which incorporate the latest technology achieve total mass filtration rates of up to 99.95%. Because the service life of a commercial-vehicle engine far outstrips that of a passenger-car engine, the filtration-efficiency requirements are greater accordingly.
The filter elements have to be replaced at the intervals specified by the vehicle manufacturer. Usual replacement intervals are, depending on the application, 40,000 to 320,000 km or two years. In most cases a service-interval indicator monitors when the service interval is reached.

The media used in commercial-vehicle air filters are comparable to filter media in passenger cars, but are characterized by higher filtration efficiency and are usually designed as surface-type filters. The separated particles thus precipitate mainly on the surface of the cellulose-based medium.

Better values than with pure cellulose-based media are obtained with special nanofiber-based filter media, in which

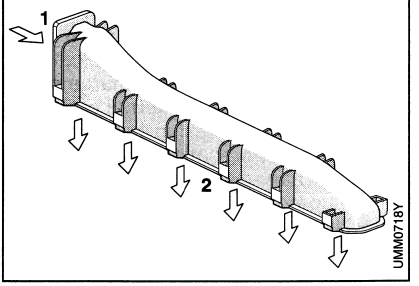
Figure 18: Commercial-vehicle air filter
1 Raw-air inlet, 2 Housing,
3 Filter element, 4 Safety element,
5 Cover, 6 Dust-discharge valve,
7 Clean-air outlet.



ultrathin, synthetic fibers with diameters of just 30 to 40 nm are applied to a relatively coarse backing layer of cellulose. The filtration-efficiency figures improved in this way enable such systems to satisfy the constantly increasing requirements of the commercial-vehicle manufacturers regarding the protection of sensors, turbochargers and other components, and by protecting these components help to comply with the ever more stringent exhaust-emissions standards.

Further development
The space allotted to the air-intake system is becoming increasingly limited, particularly due to the comprehensive exhaust-gas treatment system. For this reason, there will in future also be air filters that are flatter than round filters. This opens up the possibility of new installation positions in the vehicle, for example in front of the front wheel or underneath the driver's cab.

Figure 19: Boost manifold for commercial vehicles
1 Flange for EGR valve or auxiliary heater,
2 Flange for cylinder head.



Commercial-vehicle intake manifolds

Supercharged diesel engines require one intake manifold, usually referred to simply as a boost manifold. The air passes through the air filter to the turbocharger, and from there through the intercooler to the boost manifold. This distributes the air to the cylinders. For the most part the valve for EGR introduction (exhaust-gas recirculation) and possibly an electric auxiliary heater are flange-mounted at the inlet. The auxiliary heater helps to reach the required exhaust-gas values as quickly as possible during engine starting.

The introduced exhaust gas is mixed with the fresh air in the boost manifold and distributed as uniformly as possible to all the cylinders. The exhaust gas cooled by the EGR cooler contains, depending on the operating state and EGR rate, sulfurous acid and soot. The aluminum used for the boost manifold must be protected against the acid by for example painting. If the boost manifold is made of plastic, the plastic must be chosen so that temperature, charge-air pressure and acid do not damage the components over their service lives. The same requirements also apply to the flange gaskets.

There are both simple boost manifolds (Figure 19) which can be manufactured as a single plastic injection-molded part and also parts which due to their more complex geometry are manufactured in two parts and welded. Charge-air pressure, size of the part and tolerances must be taken into account when configuring the welding. The use of plastic also contributes to the lightweight construction of the engine.

Turbochargers and superchargers

Turbochargers and superchargers increase the air-mass flow for a given engine displacement and a given engine speed, thereby increasing the power density. These "compressors" are generally systematically divided into three different types: the mechanically driven supercharger, the exhaust-gas turbocharger, and the pressure-wave supercharger.

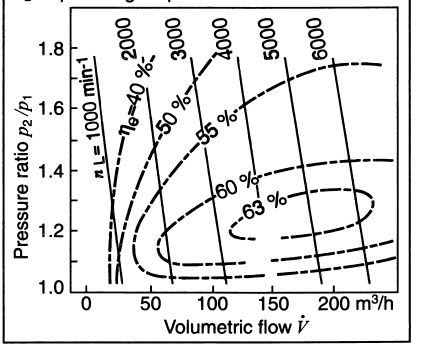
In the case of the mechanically driven supercharger, the required drive power is branched directly from the crankshaft via a belt or a gear drive from the engine, i.e. the supercharger and the engine are mechanically coupled.

In the case of the exhaust-gas turbocharger, the drive power is generated after banking up the exhaust gas through expansion of the exhaust gas, by which part of the exhaust-gas energy is still used. The turbocharger and the engine are thus purely thermodynamically coupled. In this respect, an exhaust-gas turbocharger with an electric, mechanical or hydraulic supplementary drive is a special case.

In the case of the pressure-wave supercharger, the compression power is derived from the exhaust gas, but an additional drive is required to synchronize the supercharger, meaning that the supercharger is simultaneously coupled mechanically and thermodynamically to the engine.

Figure 1: Program map of a Roots supercharger

p_1 Total pressure before compressor,
p_2 Total pressure after compressor,
n_L Supercharger speed.



Superchargers (mechanically driven)

Mechanically driven superchargers [1] fall into two categories: Those which operate according to the positive-displacement principle and those which compress air according to the flow principle and the principle of momentum. The internal-combustion engine drives the supercharger as a rule by means of a belt drive. It is driven either directly (continuous-running operation) or via a clutch.

Mechanical centrifugal supercharger

The compressor of the mechanical centrifugal supercharger operates according to the flow principle. This type of device is very efficient, providing the best ratio between unit dimensions and volumetric flow when compared with other mechanically driven superchargers. The pressure ratio that can be achieved is dependent on the peripheral velocity of the compressor impeller. The air throughputs typical of passenger-car engines give rise to small compressor-impeller dimensions. The pressure ratios required for effective engine supercharging thus require high compressor speeds.

The range of application of the mechanical centrifugal supercharger is limited technically by the high rotational speeds required and the transmittable drive power, and economically by the comparatively high costs. It is used in small numbers in performance-oriented engine concepts.

Positive-displacement supercharger

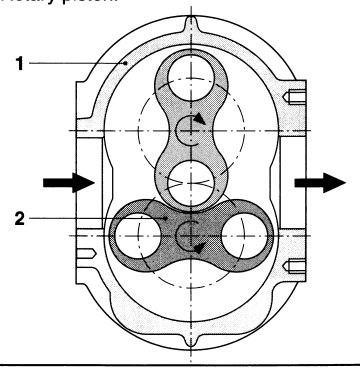
A comparatively large number of designs has been proposed for units operating according to the positive-displacement principle, but only a few have succeeded in being employed in current series applications. Positive-displacement superchargers can operate with or without internal compression. Internal-compression superchargers include the reciprocating-piston, screw-type and rotary-piston compressors. The Roots supercharger is an example of a unit without internal compression. All of these positive-displacement superchargers share certain characteristics as shown

here in the program map of a Roots supercharger (Figure 1):

- The speed curves n_L = const in the p_2/p_1-$\dot{V}$ program map are very steep, i.e. the volumetric flow $\dot{V}$ decreases only slightly as the pressure ratio p_2/p_1 increases. The drop in volumetric flow is basically determined by the efficiency of the gap seal (backflow losses). It is a function of the pressure ratio p_2/p_1 and of time, and is not influenced by rotational speed.
- The pressure ratio p_2/p_1 does not depend on the rotational speed. In other words, high pressure ratios can also be generated at low volumetric flow rates.
- The volumetric flow $\dot{V}$s independent of the pressure ratio and, roughly formulated, directly proportional to rotational speed.
- The unit retains stability throughout its operating range. The positive-displacement compressor operates at all points of the p_2/p_1-$\dot{V}$ program map as determined by supercharger dimensions.

Mechanical positive-displacement superchargers generally must be substantially larger than mechanical centrifugal superchargers in order to produce a given volumetric flow.

Figure 2: Cross-section through a Roots supercharger
1 Housing,
2 Rotary piston.



Roots supercharger

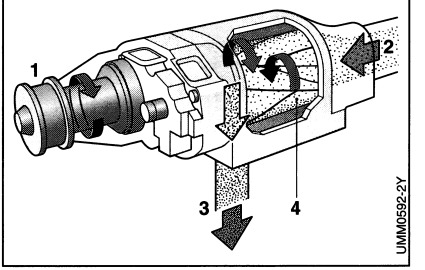
The Roots supercharger is a rotary-piston machine with contra-rotating twin- or multi-bladed rotary pistons mounted in rolling bearings and synchronized by a gear set. The rotary pistons mesh at identical speed without directly contacting each other in the housing (Figure 2). The gaps between these components essentially determine the machine's efficiency.

Roots superchargers operate without internal compression. Mufflers are generally fitted on the suction and pressure sides to limit their acoustic impact. Their system and design limits the pressure ratios that can be achieved to levels below 2. Coating the functional components has helped to improve efficiency. Current development work is focused on among other things variable-speed gearboxes on the drive side.

Screw-type supercharger

The screw-type supercharger (Figure 3) is very similar in design to the Roots supercharger, i.e. it is a twin-shaft contra-rotating rotary-piston machine. However, it differs from the Roots supercharger in that it generally operates with internal compression. It can achieve higher pressure ratios than a Roots supercharger. On the suction side (inlet) the rotation of the pistons opens a profile-gap space which fills with inducted air. As the rotors continue to contra-rotate, the profile-gap space continuously decreases in size until it reaches the outlet control edges. At this point internal compression is completed and the compressed volume is forced into the pressure connec-

Figure 3: Screw-type supercharger
1 Drive,
2 Air inlet,
3 Air outlet (compressed air),
4 Screws.



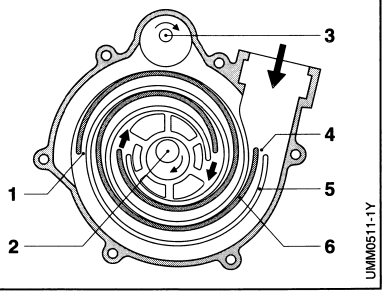
tion (outlet). In order to minimize internal leakage losses, the tolerances between the rotors and the walls must be very closely adhered to. Measures along similar lines to the Roots supercharger are considered for improving the function and efficiency of this type of supercharger.

Spiral-type supercharger
The spiral-type supercharger (Figure 4) is a compressor in which a rotor with spiral lands eccentrically describes a circular path in a housing also with spiral lands. The spiral-type supercharger operates according to the principle of a circulatory displacer element. Working chambers in sequence open for charging, close for transport, and open once again for discharge at the hub.

The G-supercharger (G-Lader) is an example of this type. Maintaining very tight tolerances and the seal has a function-determining effect. Radial sealing is provided by the narrowest gaps possible, axial sealing by touching sealing strips inserted on the end faces. These sealing strips are subject to wear and are if necessary replaced during regular vehicle servicing. Internal compression can be achieved by an appropriate configuration of spirals.

More recent developments are heading towards a simplification of the design and the integration of a switchable clutch.

Figure 4: Cross-section through a spiral-type supercharger
1 Air inlet into second working chamber,
2 Drive shaft,
3 Displacer element,
4 Air inlet into primary working chamber,
5 Housing,
6 Displacer element.



Pressure-wave superchargers

Design and operating principle
The pressure-wave supercharger (Figure 5) is a gas-dynamic machine whose core component is a rotor with open channels arranged coaxially around its circumference ("cell wheel" or "rotor") [2, 3]. The inlet and outlet openings for fresh air and exhaust gas pressurize the channels via the rotor end faces. The fresh air is compressed in the rotor channels by way of gas-dynamic processes. Fresh gas and exhaust gas briefly come into direct contact with each other in the process. Essential to the function is the physical fact that the gas-dynamic compression process takes place in a much shorter period of time than the mixing of the two gas flows.

The operating principle of the pressure-wave supercharger is based on the fact that a pressure wave is reflected at an open end of a line as a vacuum wave and at a closed end as a pressure wave; this also applies accordingly to the reflection of a vacuum wave. To control and maintain the pressure-wave process, the channel openings pass over "open ends" and "closed ends". The cell rotor must rotate for this purpose. The drive power serves merely to cover the rotor bearing and ventilation losses and to accelerate the rotor in the event of a sudden load variation.

Energy is exchanged in the channels at the velocity of sound. Owing to the principles involved, the pressure-wave supercharger responds very quickly to changes of state. with the actual response times of the supercharger being determined by the charging processes in the air and exhaust pipes. The velocity of sound is, as well as the physical characteristics, a function of temperature, meaning that essentially it is dependent on engine torque, but not engine speed.

Comprex supercharger
If a constant gear ratio is employed between engine and cell rotor, as is achieved with the belt-driven Comprex supercharger, the pressure-wave process is, on account of this boundary condition, optimally tailored to only a single operating point. To get around this disadvantage, appropriately designed "pockets" can be incorporated in the gas routing of the end housings. These achieve high efficiency levels extending through a relatively wide range of engine operating conditions and provide a good overall boost curve.

The Comprex supercharger's rotor is over-mounted and is provided with permanent grease lubrication, with the bearing located on the unit's air side. An integral bypass valve in the supercharger regulates boost pressure according to demand.

The Comprex supercharger did not meet with the desired response from the European vehicle manufacturers.

Hyprex supercharger
A further development of the Comprex is the Hyprex, which provides an effective supercharging process particularly for small-capacity gasoline engines. However, its high exhaust-backpressure sensitivity has hitherto come out against series application.

The Hyprex supercharger's rotor is driven independently of engine speed by an electric motor, thanks to which the rotor speed can be better adapted to the engine's operating state. In addition to other modifications which improve cold starting, asymmetrical arrangement of the cells has further improved acoustics. A variable gas-pocket inflow provides for improved efficiency in the lower engine speed range with a resulting increase in boost pressure. Comprehensive application of a Hyprex supercharger requires the use of modern electronic engine management.

Exhaust-gas turbochargers

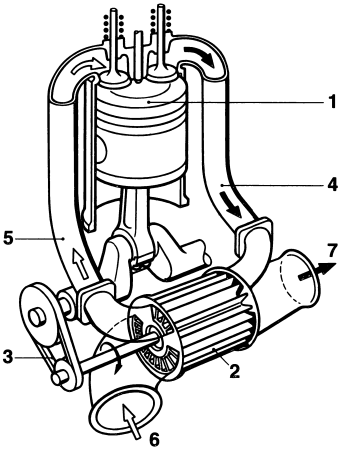
Of all the supercharger units in use, the exhaust-gas turbocharger is by far the most widespread. While nowadays virtually 100 % of passenger-car and commercial-vehicle diesel engines are turbocharged, the proportion of supercharged gasoline engines will continue to increase dramatically in the next few years. Here, too, the exhaust-gas turbocharger will be the dominant supercharger.

Design and operating principle

The exhaust-gas turbocharger consists of two turbo elements: a turbine and a compressor, where the impellers for the two elements are mounted on a common shaft (Figure 6 and Figure 7). The operating concept of a turbine engine is based physically on the principle of momentum. The turbine converts some of the enthalpy contained in the exhaust gas into mechanical energy to drive the compressor. The compressor draws in fresh air through the air filter and compresses it.

The exhaust-gas turbocharger is coupled only thermodynamically, not mechanically,

Figure 5: Pressure-wave supercharger
1 Engine, 2 Cell rotor, 3 Belt drive,
4 High-pressure exhaust gas,
5 High-pressure air, 6 Low-pressure air inlet,
7 Low-pressure gas outlet.



to the engine. Turbocharger speed does not depend on engine speed, but is rather a function of the balance of drive energy between the turbine and the compressor. The turbine generates the compressor output and the output which is essentially used up in the bearings and ultimately dissipated as heat (mechanical losses).

Application

The exhaust-gas turbocharger in its present form can trace its roots back to the work of Alfred Büchi (1905) [2], who had already recognized the potential of combining supercharging and valve overlap to scavenge residual exhaust gas (1915). Exhaust-gas turbochargers have traditionally been used for the purpose of supercharging large diesel engines, originally being used primarily for truck, marine and locomotive power plants as well as for agricultural- and construction-machinery applications.

Use in diesel vehicles

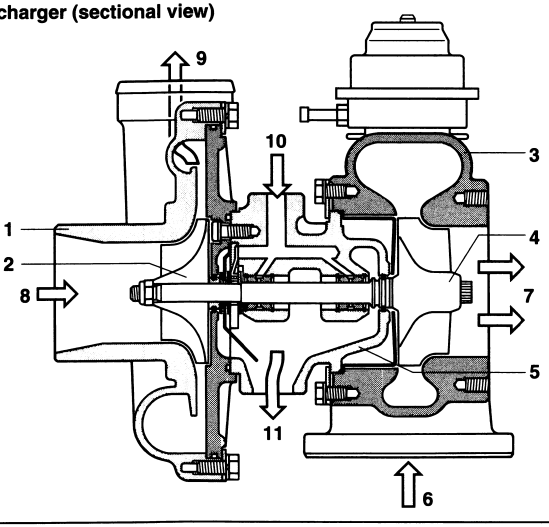
The mid-1970s saw the advent of the first production diesel engines supercharged by exhaust-gas turbochargers in passenger cars. The introduction of a "wastegate" for controlling boost pressure finally facilitated a torque-oriented design and thereby improved driveability. Further increases in the performance of passenger-car diesel engines have been achieved through the introduction of direct injection (1987) and by exhaust-gas turbocharging with variable turbine geometry (1996) or two-stage exhaust-gas turbocharging (2004). These developments have been reflected in a marked increased in the market share for diesel engines in Europe. Today in Europe all passenger-car and commercial-vehicle diesel engines are equipped with exhaust-gas turbocharging and intercooling (charge-air cooling).

Use in passenger cars with gasoline engines

Exhaust-gas turbocharging of gasoline engines was originally the reserve of only high-output sports engines for increased performance and, because of inadequate driveability ("turbo lag"), was rarely encountered on the market. In the meantime, gasoline-engine supercharging has become an integral part of engine development, predominantly in connection with small to medium-sized engines. In addition to the improvement in efficiency, one of the main objectives is to avoid the increases in the number of engine cylinders (e.g. from four to six cylinders), thus positively influencing package space and fuel consumption.

Figure 6: Exhaust-gas turbocharger (sectional view)

1 Compressor housing,
2 Compressor impeller,
3 Turbine housing,
4 Turbine wheel,
5 Bearing housing,
6 Incoming exhaust gas,
7 Outgoing exhaust gas,
8 Atmospheric fresh air,
9 Precompressed fresh air,
10 Oil supply,
11 Oil return.



In contrast to diesel engines, mechanical supercharging was also employed and today to a lesser extent continues to be employed on gasoline engines for marketing reasons and on account of the excellent transient build-up of boost pressure. With the introduction of direct injection in the gasoline engine the exhaust-gas turbocharger was almost on a par with mechanical superchargers thanks to the use of scavenging.

In addition, a combination of mechanical supercharging and exhaust-gas turbocharging (combination supercharging) was in the past employed to provide small-capacity gasoline engines with high rated output and high torque at a comparatively low specific speed. However, the current trend is towards the use of electrically driven supplementary compressors instead of combination supercharging.

While the exhaust-gas turbocharger with variable turbine geometry is today the standard supercharger used in diesel engines, the high exhaust-gas temperature and the costs associated with this technology means that up to now it has only been used on gasoline engines in niche applications.

With regard to the statutory exhaust-emission and consumption guidelines and the existing performance figures required by customers, the importance of exhaust-gas turbocharging to the reduced-capacity and reduced-cylinder engines currently in development ("downsizing") will increase further. Today, the number of turbocharged gasoline engines is rising dramatically and in the next few years a further dramatic growth of the market segment is expected.

Exhaust-gas turbocharger design

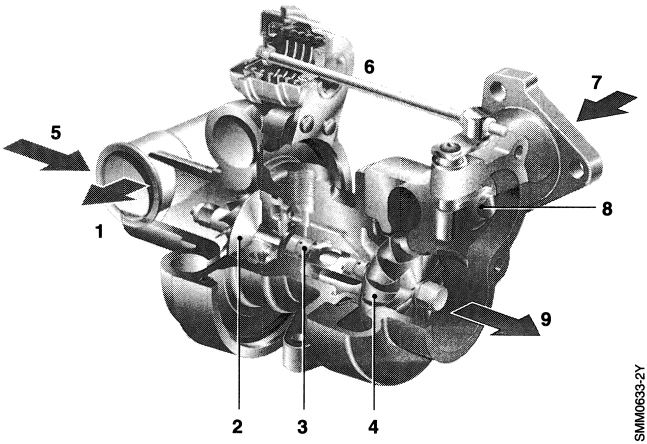
The exhaust-gas turbocharger essentially consists of three assemblies: the bearing housing, the compressor, and the turbine (Figures 6 and 7). Depending on the design, there may be a fourth assembly: a boost-pressure control facility.

Bearing housing

The bearing housing accommodates the bearings and the elements for sealing the shaft. State-of-the-art designs usually feature a specially developed plain bearing both in the radial as well as the axial (thrust) bearing assembly. The radial bearings are designed either as rotating double plain bushings or as stationary plain-bearing bushings. The thrust bearing in conventional designs is made up of a multiple-spline surface bushing that is subject to load from both sides.

Figure 7: Exhaust-gas turbocharger with wastegate (view)

1 Compressed-air outlet,
2 Compressor impeller,
3 Shaft,
4 Exhaust-gas turbine,
5 Fresh-air inlet,
6 Linkage,
7 Exhaust-gas mass-flow intake,
8 Wastegate (in turbine inlet),
9 Outlet, exhaust-gas mass flow.



The lubricating oil is supplied by connecting the turbocharger to the engine's oil circuit. The oil outlet is connected directly to the oil pan in the crankcase. Today, this type of bearing assembly is used to control rotational speeds of more than 300,000 rpm reliably.

Rolling-bearing exhaust-gas turbocharger rotors are expected to deliver consumption-related efficiency advantages which provide for a better transient build-up of boost pressure and higher boost pressures in the part-load range. Rolling-bearing rotors are now being used in series production for the first time. For deeper market penetration, however, further developments with regard to price, acoustics and service life are necessary.

To seal off the oil chamber to the exterior and to minimize the entry of charge air ("blow-by") and exhaust gas to the charger interior, the shaft is equipped at the casing openings with piston rings which are tensioned in the bearing housing and with the shaft groove form a simple labyrinth. The state-of-the-art sealing technology used today limits the permissible installation location for the exhaust-gas turbocharger to a comparatively small range of inclination. Touching seals would be useful in this respect, but have not been able to be produced up to now at reasonable cost owing to the high relative speeds between shaft and housing.

No additional cooling measures are necessary to maintain efficient operation of the bearing assemblies under suitable ambient conditions in the engine compartment, i.e. at exhaust-gas temperatures up to approximately 820 °C. The relevant temperatures can generally be maintained below critical levels using devices such as a heat shield, and by thermally isolating the hot turbine housing, supplemented by incorporating suitable design elements in the bearing housing itself. Water-cooled bearing housings are employed for higher temperatures, for example, on gasoline engines operating at up to 1,050 °C exhaust-gas temperature or in certain diesel-engine applications.

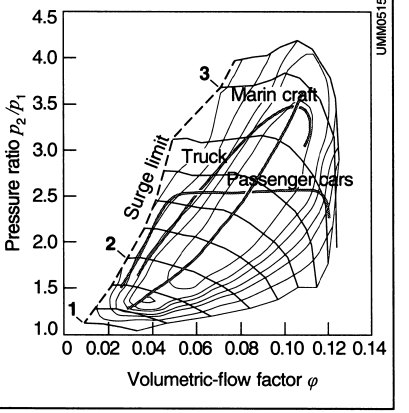
Compressor

The compressor assembly in passenger car and commercial-vehicle applications consists of an axial-inflow and radial-outflow impeller (centrifugal design) and a generally guide-vane-free compressor housing made of cast aluminum. The flow is further slowed down in the housing, in the course of which some of the pressure build-up occurs. Having the smallest possible contour gap between the impeller and the housing is crucially important to efficiency.

The impeller is mass-produced in a special casting process from an aluminum alloy. Today, however, impellers milled from an aluminum forging alloy are increasingly being used. In the event of more rigorous service-life requirements, particularly in the commercial-vehicle sector with critical application profiles, impellers milled from a titanium alloy are also used, which is a comparatively cost-intensive measure. For engines with low-pressure-side exhaust-gas recirculation, where the compressor has to deliver medium containing exhaust gas, the compressor impeller is provided with a coating. This protects the

Figure 8: Compression graph with typical engine operation curves valid for all displacements

1 Peripheral velocity u = 150 m/s,
2 Peripheral velocity u = 300 m/s,
3 Peripheral velocity u = 450 m/s.
u Peripheral velocity,
p_1 Total pressure at compressor inlet,
p_2 Total pressure at compressor outlet.
The volumetric-flow factor is a dimensionless volumetric flow.



impeller against erosion by particulate and droplet impingement.

Although the compressor impeller contributes much less than the turbine wheel to the rotor's mass moment of inertia, a value which is as low as possible is also important on the cold side. Because of the heightened requirements with regard to performance and service life, attempts to mass-produce plastic compressor impellers have hitherto failed.

The characteristic of a compressor is described by a program map (Figure 8). In contrast to a positive-displacement supercharger, the program map of a flow compressor features ranges in which stable operation is not physically possible. By implementing appropriate measures, the effective range as well as the speed and efficiency characteristics of the compressor can be adapted to the required boost-pressure curve. The effective range of the compressor is determined on the "left" side of program map (i.e. maximum achievable pressure ratio for a given throughput) by the surge limit. On the "right" side (maximum possible throughput, limited by the attainment of the velocity of sound in the narrowest cross-section) the choke limit restricts the operating range. The surge limit is defined as the transition from the stable to the unstable operating range. Unstable means that, triggered by a severance in flow – generally at the compressor-impeller inlet – the air-mass flow is periodically interrupted and re-established to produce a pumping effect. Among other factors, the extent of the surge limit is also governed by the design of the intake line. The surge limit is therefore not a compressor property, but rather a system property. The choke limit, identifiable by the range of steeply descending speed curves in the program map, is determined based on the free inlet cross-section of the compressor impeller and therefore across the impeller diameter.

In view of the proportionality of the engine speed and volumetric air flow of the compressor, it is easy to deduce that compressors for turbocharging gasoline engines must have a substantially larger effective map width than, for instance, compressors which supply large diesel engines with air. The dimensionless compressor program map in Figure 8 shows the characteristic air requirements along the WOT for passenger cars, commercial vehicles, and large heavy-duty engines.

Precise aerodynamic adaptation of the compressor to the respective requirements is effected by modifying the component geometry, where impeller and housing must be viewed as a single component and strength criteria and natural-oscillation responses must be taken into account in the design of the impeller. A diffuser with vanes may indeed increase efficiency at the design point, but with an otherwise identical geometry gives rise, apart from a blocking effect, to a downturn in efficiency when the design point is deviated from. For this reason only compressors without guide vanes are used for passenger-car turbocharging in view of the comparatively large mass-flow spread here.

Recirculation channels integrated into the housing ("blow-off" by returning some of the delivered mass flow to the compressor inlet) or a special form of the compressor inlet channel (preliminary volumes) are also used to extend and stabilize the program map.

The program map can be extended in the surge-limit range by components such as variable inlet or discharge guide vanes. This means that the compressor can be operated at an operating point at which the surge limit would already be exceeded. Apart from a simple preliminary diffuser, whose vane angles twist into a specific position as the mass flow increases (flow forces) and are released again as the mass flow decreases, variable compressors have up to now not been series-produced.

Compressors for turbocharging gasoline engines have until now been equipped with a blow-off valve, whose task it is to prevent compressor surge when load is quickly removed (i.e. throttle valve closes). The blow-off valve, which was originally actuated pneumatically but is now mainly actuated electrically, for this purpose creates a bypass between the compressor outlet and inlet. In this way, the compressor pumps briefly as a blower in the circuit. Today it is possible, above all by introducing electrical actuation of the wastegate, to prevent surge without the use of a blow-off valve and guarantee reliable operation.

Turbine

The turbine consists of a diffuser and a wheel. The diffuser of the turbocharger turbine is integrated in the flow housing (spiral housing) and for passenger-car applications, if the turbocharger is one without variable turbine geometry, is designed without vanes. Fixed-geometry diffusers with vanes are frequently used in exhaust-gas turbochargers for large engines for better fine graduation of the turbine throughput. The flow is accelerated in the diffuser and distributed as evenly as possible to the turbine wheel.

The turbine wheel is normally designed as a centripetal turbine, i.e. radial inflow and axial outflow. Semiaxial wheels are used on account of their favorable mass moment of inertia. These have diagonal inflow and axial outflow. Axial wheels with axial inflow and axial outflow are only used in heavy-duty turbine applications. The type of construction determines the achievable pressure ratios, efficiencies, and other characteristics. Generally speaking: The greater the mass flow, the more advantages the axial machine offers and vice versa.

The arrangement of the exhaust-gas pipe plays an important role in supercharging. Classically, a distinction is made between pulse supercharging and constant-pressure supercharging. In the case of pulse supercharging, the exhaust-gas pipes are routed separately to the turbine housing. Those cylinders whose exhaust pulses influence each other minimally are brought together on the exhaust-gas side. This means that the scavenging processes of the cylinders that are brought together are subject to less hindrance. The turbine housing is designed in such a way that the separation of channels is maintained as far as possible up to the turbine-wheel inlet. The channel cross-section is the size of the exhaust cross-section. The exhaust pulses are essentially transmitted as pressure waves to the turbine such that the exhaust-gas energy is applied at the turbine inlet in highly pulsating form. This effect can be utilized particularly well at low engine speeds where pressure surges can develop effectively due to the long time interval from one exhaust pulse to the other.

In the case of constant-pressure supercharging, the exhaust discharges are directed in a common, comparatively large-volume collector to the turbine housing. This process smooths out the individual exhaust pulses to a large extent.

With constant-pressure supercharging the more uniform enthalpy over time can be processed with greater efficiency. Pulse supercharging, on the other hand, offers advantages for the engine's part-load and accelerating performance.

The design of the turbine housing differs significantly, depending on whether constant-pressure or pulse supercharging is used. Pulse supercharging for the most part is used in commercial-vehicle applications. The turbine housing is designed as double-flow housing in which the two flows join only just before the wheel inlet.

Constant-pressure supercharging is used on high-speed engines such as the diesel engines in passenger vehicles. The turbine housings are single-flow ("single-scroll") in design and partially integrated into the exhaust manifold as a single component. This provides for a particularly compact and streamlined component geometry.

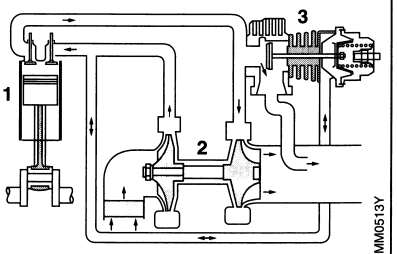
Gasoline engines utilize both constant-pressure and pulse supercharging. The exhaust systems are configured accordingly and the housings are single-flow or twin-flow ("twin-scroll" or "dual volute") in design. Partially manifold-integrated turbine housings are also used in exhaust-gas turbochargers for gasoline engines.

To withstand the high temperatures and high loads involved, the turbine wheels are made of a material containing high levels of nickel using the lost-wax process and joined to the steel rotor shafts by means of friction, laser or electron-beam welding. Because of its outstanding high-temperature properties, the material titanium aluminide, an intermetallic phase, is being discussed as a potential material for the turbine wheel. Thanks to its low density, it reduces significantly the mass moment of inertia and allows the wheel to be fashioned with a raised wheel back without negatively impacting on the turbocharger's run-up performance. On account of its brittleness at room temperature and problematic castability, development and production issues must still be resolved in this regard.

The turbine housings made of high-alloy spheroidal or spherulitic cast iron, depending on the specific application temperature, are produced in an open sand-casting process. Alloyed cast steels, normally cast in thin walls using the lost-wax process, are used in high-temperature applications (up to 1,050 °C). Modern low-pressure casting processes are also used to manufacture thin-walled components (weight, costs, thermal inertia). Twin-scroll housings, on account of the filigree geometry of the partition wall, require cast-steel materials. Turbine housings built up from welded sheet-metal parts in single-walled or twin-walled form have also been produced in AGI versions (AGI: Air-Gap-Insulated). AGI components reduce wall heat losses and have a low thermal inertia.

So far built-up turbine housings have only been introduced in series production to a limited extent. However, these could help to meet future emission regulations in that they enable the catalytic converter to be heated up more quickly after cold starting.

Figure 9: Boost-pressure control via exhaust-side boost-pressure control valve (wastegate)
1 Engine,
2 Exhaust-gas turbocharger,
3 Boost-pressure control valve (wastegate).



Boost-pressure control
In view of the large rotational speed spread of passenger-car engines, a boost-pressure control facility is indispensable to maintain the maximum permissible boost pressure for an acceptable design torque. Standard practice presently favors turbine output control on the exhaust-gas side.

Bypass control
A simple and in gasoline and commercial-vehicle engines widely used method is that of bypass control, in which generally some of the exhaust-gas flow is routed through a flap valve ("wastegate") around the turbine (Figure 7 and Figure 9). The flap valve is actuated by a pneumatic unit or an electric actuator. The pneumatic unit can take the form of an overpressure unit (supplied by the boost pressure itself) or a vacuum unit (vacuum supply from the vehicle system). If the boost pressure is used, the wastegate cannot however be actuated independently of the engine operating state. In most cases, the control pressures are regulated by using clocked-pulse valves. Because of the thermal load, the actuators are generally mounted on the compressor side and connected via a control rack to the wastegate lever.

More recent developments are clearly heading towards the use of electric actuators that can adjust the control characteristics faster and with greater precision, and towards actuating the wastegate independently of the engine operating state. This is advantageous when it comes to compliance with emission and consumption legislation. In addition, with electric actuators it is possible to create locking forces in a tenable space which ensure that the wastegate is firmly seated in all operating states, which in turn provides for a rapid buildup of boost pressure.

Variable turbine geometry
Compared to bypass control, variable turbine geometry (VTG) offers a much more efficient option for adapting the supercharger within the overall program map. Today it is state of the art in passenger-car diesel engines and will in future also be introduced in commercial-vehicle and gasoline-engine applications.

The entire mass exhaust-gas flow is routed here via the turbine, providing specific benefits in exploiting the available energy. By varying the effective turbine-inlet cross-section, the turbine's resistance to flow is adjusted corresponding to the required boost pressure level.

Of all the potential designs, adjustable blades have achieved general acceptance, as they combine a wide control range with high efficiency levels (Figure 10). The blade angle can be easily adjusted by means of a swiveling movement. The blades, in turn, are set to the required position using adjusting cams, or directly via control levers attached to the individual blades. All blades engage in an adjusting ring which can be turned by means of a lever. The lever is actuated by pneumatic or electric actuators. Where a pneumatic unit is fitted, today a position sensor is increasingly integrated in the unit, and the signal from this sensor ensures that the VTG blade position is clearly identified for engine management.

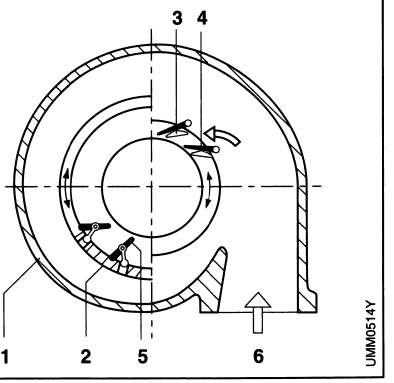
Where originally symmetrical blades with straight median lines were installed, today profiled blades with pronounced pressure and suction sides and curved median lines are used. This helps to improve efficiency above all in the closed position, i.e. when accelerating and in the part-load range. Maintaining small gaps in the guide-vane area is of vital importance to engine performance.

Owing to the high exhaust-gas temperatures of a gasoline engine, the operationally reliable and durable implementation of an adjusting-blade exhaust-gas turbocharger for gasoline engines involves very intensive thermomechanical matching and the use of highly developed materials. For this reason an adjusting-blade VTG was for a long time only available in just one gasoline-engine application of one sports-car manufacturer. In 2016, however, this technology also went into series production in mid-size passenger cars with gasoline engines.

In addition to adjusting-blade VTG, a Japanese manufacturer is employing a comparatively simple variability in the field of the turbine spiral, but which fails to attain the performance levels of adjusting-blade VTG.

In addition to the adjustment mechanisms already introduced in series on gasoline engines, there is the further promising operating concept of the sliding-sleeve supercharger. In this concept, in addition to the comparatively high efficiency, the number of parts that move in the hot-gas flow is smaller than in the adjusting-blade concept and an internal bypass can be comparatively easily opened to bypass the turbine with just one actuator in addition to actuating the variability.

Figure 10: Variable turbine geometry (schematic)
1 Turbine housing, 2 Adjusting ring, 3 Control cam, 4 Adjusting blade, 5 Adjusting blade with control lever, 6 Air inlet.



Exhaust-gas turbochargers with supply of additional energy

Modern exhaust-gas turbochargers are characterized by high efficiency levels and a low mass moment of inertia on the part of the rotor such that a very good transient build-up of boost pressure is achieved, particularly in conjunction with measures on the engine, such as for instance "scavenging" [4]. In the past various solutions have been proposed in order in states with little available exhaust-gas energy to supply additional mechanical, hydraulic or electrical power to the shaft for the purpose of accelerating the rotor more quickly.

For example, the incorporation of a gear on the turbocharger shaft allows mechanical energy to be supplied via a gearbox and a switchable mechanical connection to the crankshaft.

A further approach provides for installing on the rotor between the bearings a

Pelton turbine wheel which is pressurized with high-pressure oil (approx. 100 bar) from the engine circuit or separately with hydraulic fluid.

In a further approach it is proposed to install on the turbocharger shaft, for instance between the bearings, an electric motor which is supplied via power electronics by the vehicle electrical system ("electrically assisted exhaust-gas turbocharger"). The electrically assisted turbocharger is able in certain operating phases to function as a generator and thereby return electrical energy to the vehicle electrical system, but a wastegate cannot be dispensed with.

What all the approaches have in common is the fact that in operating phases in which no additional energy is supplied the components installed on the turbocharger shaft must be carried along and thereby cause losses. They are also subject to considerable thermomechanical load. The potential of these approaches must further be considered against the background that additional energy, partially isolated from the engine's operating state, can indeed be supplied, but the program-map limits of the compressor used cannot thereby be extended or shifted. In addition, all the solutions described involve a significant increase in expenditure on equipment and thereby in the costs. None of these approaches has been introduced into series production.

Complex supercharging systems

With a parallel or multistage connection of supercharger units the power limits can be significantly extended compared with supercharging with just one single-stage supercharger unit. In part complex valves for control, for operating-point-dependent cut-in or for defined distribution of the exhaust-gas mass flow are required. The objective here is to improve the air supply on both a stationary and a non-stationary basis and at the same time to improve the specific consumption of the engine.

Bi-turbo supercharging

In the case of bi-turbo supercharging, the air supply, instead of using one large exhaust-gas turbocharger, is distributed to two or more smaller exhaust-gas turbochargers connected in parallel. These are generally permanently assigned to specific cylinders on the exhaust-gas side, for example the bank of a V-engine. Each turbocharger pressurizes the engine with air over the full rotational-speed range. The advantage of this arrangement lies primarily in the rapid transient build-up of boost pressure and in the comparatively compact routing of pipework. The turbochargers operating in parallel must be identical in terms of their control engineering.

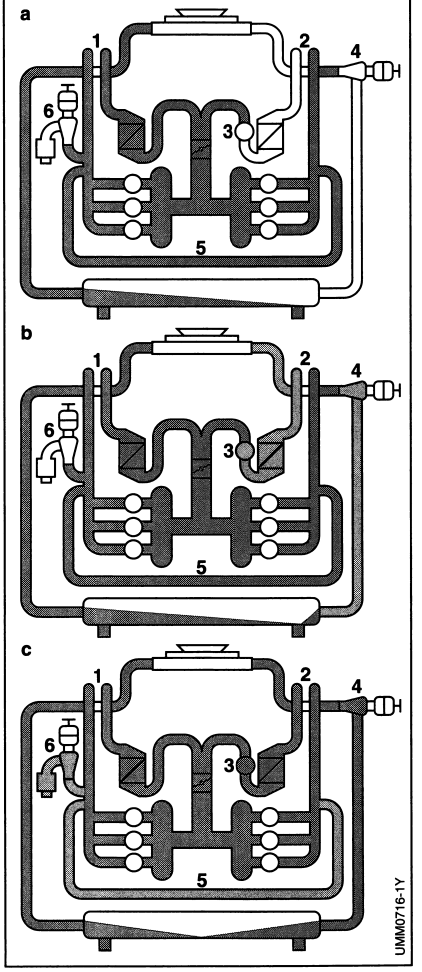
Sequential supercharging

Sequential supercharging is predominantly used in ship propulsion systems or generator drives. However, because of the high levels of power it can deliver, this arrangement variant is also used in passenger cars. In the case of passenger-car sequential supercharging, as engine load and speed increase, an additional turbocharger is cut in to the basic supercharging process (Figure 11). Thus, in comparison with mono-supercharging (i.e. a single charger which is geared to the rated output), two supercharging optima are achieved.

The sequential-supercharging system contains, in addition to the turbochargers, valves and sensors which facilitate a smooth switchover from one- to two-charger operation and back from two- to one-charger operation. When the engine is in the lower speed range, only the basic

turbocharger is active, supplying all the cylinders with air. When a certain speed is reached, the second turbocharger is run up and cut in. Both turbochargers are now operating in parallel to pressurize all the cylinders. Sharing the air supply between two turbochargers helps to deliver a highly responsive torque curve and high rated output. Furthermore, the significantly lower mass moment of inertia of the smaller turbocharger shafts gives rise to a favorable transient response.

Figure 11: Sequential supercharging
a) Run-up in one-charger operation,
b) Controlled one-charger operation,
c) Two-charger operation.
1 Turbocharger 1, 2 Turbocharger 2,
3 Compressor cut-in valve,
4 Turbine cut-in valve,
5 Transverse pipe, 6 Bypass valve.



Even an integral design for the two turbochargers including the switchover valves is possible, but complex. Controlling the switchover process (by means of control engineering) is comparatively complex, but technically well within the realms of feasibility with the methods of modern electronic engine management.

Sequential supercharging is used in both gasoline and diesel engines [5].

Dual-stage controlled supercharging

In contrast to sequential supercharging, with dual-stage controlled supercharging the turbochargers are connected in such a way as to effect serial operation in certain operating phases (Figure 12). The advantage of this supercharging system over the single-stage method lies in the increase in rated output with a simultaneous improvement in the stationary torque at low speeds and in the engine's accelerating performance thanks to the rapid build-up of boost pressure. This extremely powerful system features the use of two turbochargers of significantly different sizes. There is no dip in torque when this system is appropriately tuned and controlled.

The entire inducted fresh-air mass flow is precompressed by the low-pressure stage. The charge air is further compressed in the high-pressure stage. As a result of precompression, the relatively small high-pressure compressor works at a higher pressure level so it is able to generate the required air-mass flow. Performance can be further improved by cooling the charge air after the first compression stage.

Particular importance is attached to the sealing and the control quality of the turbine bypass valve in the high-pressure stage. By way of this valve, depending on the operating point, either the entire exhaust-gas mass flow is expanded in two stages in the high-pressure turbine and then in the low-pressure turbine, or some of the exhaust-gas mass flow is diverted past the

high-pressure turbine and delivered directly to the low-pressure turbine. The thermomechanical and vibrational demands placed on the valve system are very high. Routing flow between the compressors and the turbines with minimal pressure losses is also crucially important.

The bypass valves are closed in the engine's low speed range, i.e. with small exhaust-gas mass flow rates. The exhaust-gas mass flow expands primarily through the small high-pressure turbine. This results in a very good transient build-up of boost pressure. As engine speed increases, the high-pressure turbine bypass valve routes an increasing amount of the exhaust-gas mass flow directly to the low-pressure turbine until finally the low-pressure turbine too is relieved by a further wastegate. Thus for small throughputs the charge air is compressed predominantly in the high-pressure stage; however, as throughputs increase it is compressed increasingly in the low-pressure stage. When a specific engine speed is reached, the high-pressure compressor in spite of precompression is no longer able to throughput the air-mass flow and would in this case act as a throttle. To prevent this, the high-pressure compressor bypass valve opens a channel directly from the low-pressure compressor to the intercooler.

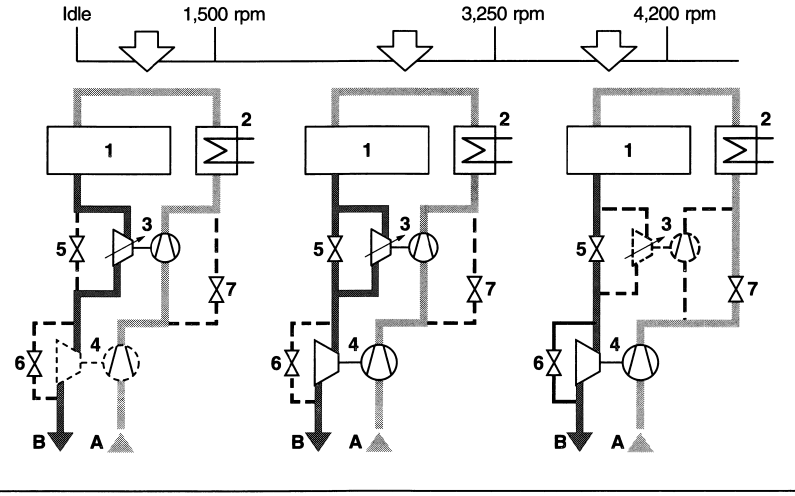
In this operating mode dual-stage controlled supercharging enables infinitely variable adaptation of turbocharging to the requirements of engine operation.

The compressor bypass valve can be designed as a (passive) non-return valve or as an actively actuated flap with actuator. With the latter variant the pressure loss in the compressor bypass can be reduced, but this involves increased control-engineering complexity.

Both fixed-geometry turbines and VTG turbines, in each case with an additional bypass valve, are used in the high-pressure stage. Either a wastegate turbine or a variable-geometry turbine (usually without an additional bypass) can be used as the low-pressure turbine.

Dual-stage controlled supercharging was introduced as standard in passen-

Figure 12: Schematic diagram of dual-stage controlled supercharging
1 Engine, 2 Intercooler, 3 Turbocharger high-pressure stage,
4 Turbocharger low-pressure stage, 5 Turbine bypass valve,
6 Wastegate, 7 Compressor bypass valve.
A Intake air, B Exhaust gas.



ger-car diesel engines in 2004. It is technically possible to implement this approach in gasoline engines too.

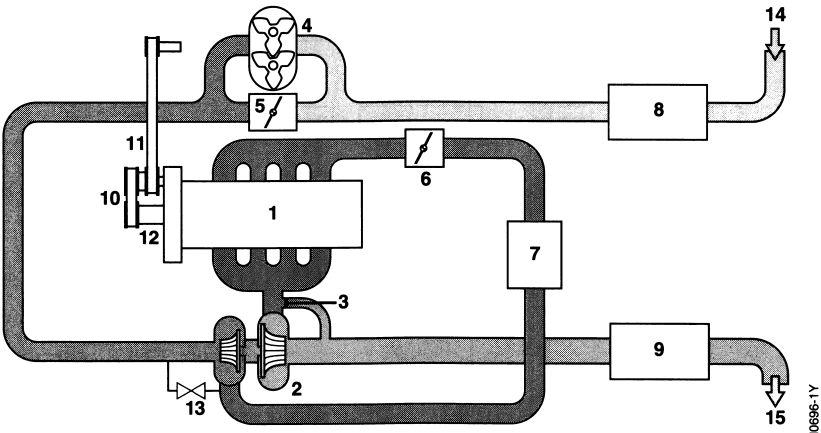
Combination supercharging
Combination supercharging combines a mechanically driven supercharger unit with an exhaust-gas turbocharger (Figure 13). This powerful supercharging system was used for the first time in a passenger car in 1985 (Lancia Delta). It can be used in comparatively small-capacity gasoline engines with high rated output and high torque already at comparatively low speeds and high dynamic response.

In the lower engine-speed range in which little exhaust-gas energy is available to drive the turbocharger turbine, the air supply is practically managed by the mechanically driven supercharger on its own. Because the exhaust-gas turbocharger is designed for the mean-speed and rated-output ranges, this results in a comparatively high-flow-rate charger with correspondingly high levels of efficiency. In a certain operating range the system operates as a series connection of two supercharger units, i.e. the total pressure ratio is the product of the individual pressure ratios.

The mechanically driven supercharger can be positioned in the flow direction before or after the turbocharger compressor. When the engine is in the upper speed range and there is therefore an adequate supply of exhaust-gas energy, supercharging is performed solely by the turbocharger and the mechanical supercharger is bypassed. Because friction losses are branched off as drive power from the crankshaft even when the mechanical supercharger is idling, disengagement of the mechanical supercharger is the most efficient solution.

Electrically driven compressor in series with an exhaust-gas turbocharger
One possible solution for obtaining the benefits of exhaust-gas turbocharging in quasi-stationary operation and in the process improving the response of a turbocharger designed for high rated output is the series connection of an exhaust-gas turbocharger with an electrically driven flow compressor ("booster", Figure 14). The advantages of this solution over the electrically assisted

Figure 13: Combination supercharging
1 Internal-combustion engine, 2 Turbocharger, 3 Wastegate, 4 Compressor (mechanical supercharger), 5 Control flap, 6 Throttle valve, 7 Intercooler, 8 Air filter, 9 Catalytic converter, 10 Belt drive for accessories with clutch, 11 Belt drive for compressor, 12 Crankshaft, 13 Blow-off valve, 14 Fresh air, 15 Exhaust gas.



exhaust-gas turbocharger are that the effective program-map range is extended by the series connection of two flow compressors and the booster can be placed in suitable locations (thermomechanical load) in the engine compartment. The booster can be positioned in the flow direction before or after the turbocharger. A bypass allows the booster to be bypassed when it is not in operation.

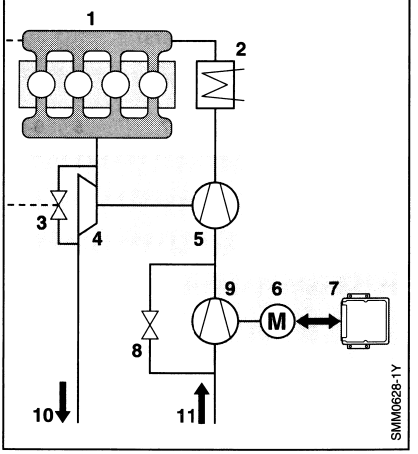
One potential mode of operation for this system, especially with regard to the electrical energy available in conventional vehicles, is operation of the electrically driven stage exclusively in transient operating phases in the engine's lower speed range.

For the system to operate effectively, the booster must reach its target pressure ratio within the shortest possible time. When the acceleration of the shaft is taken into consideration, this produces a brief booster power peak. Especially in view of the prevailing vehicle system voltage of 12 V, this requires considerable, cost-intensive additional equipment in the vehicle electrical system. The quality of the boost process and thus the effectiveness and the performance of this supercharging system are crucially dependent on the provision of electrical energy from the vehicle electrical system.

This supercharging system has so far only been implemented in some luxury-class vehicles with 48 V electrical systems. This technology could gain further currency if the increased vehicle-system voltage were to be comprehensively introduced. Active cooling of the electric motor can also deliver continuous operating capacity on the part of the supplementary compressor and thereby improve the engine's torque curve in the low-rpm range.

Figure 14: Electrically driven compressor in series with an exhaust-gas turbocharger

1 Internal-combustion engine,
2 Intercooler,
3 Bypass valve for turbocharger,
4 Turbocharger turbine,
5 Turbocharger compressor,
6 Electric motor,
7 Triggering electronics for electric motor,
8 Bypass valve for eBooster,
9 eBooster compressor,
10 Exhaust gas,
11 Intake air.



References

[1] H. Hiereth, P. Prenninger: Aufladung der Verbrennungskraftmaschine. Springer-Verlag, 2003.
[2] H. Pucher, K. Zinner: Aufladung von Verbrennungsmotoren: Grundlagen, Berechnungen, Ausführungen. 4th Edition, Verlag Springer Vieweg, 2012.
[3] L. Flückiger, S. Tafel, P. Spring: Hochaufladung mit Druckwellenlader für Ottomotoren. MTZ 12/2006, No. 67.
[4] R. v. Basshuysen (Editor): Ottomotor mit Direkteinspritzung. 3rd Edition, Verlag Springer Vieweg, 2013.
[5] J. Portalier, J.C. Blanc, F. Garnier, N. Schorn, H. Kindl: Twin Turbo Boosting System Design for the New Generation of PSA 2.2 liter HDI Diesel Engine. 11th Supercharging Conference, Dresden, 2006.
[6] M. Mayer: Abgasturbolader – Sinnvolle Nutzung der Abgasenergie. 6th Edition, Verlag Moderne Industrie, 2011.
[7] G. Hack, G.-I. Langkabel: Turbo- und Kompressormotoren – Entwicklung und Technik. 3rd Edition, Motorbuch-Verlag, 2003.

Exhaust-gas system

Purpose and design

In compliance with legal requirements, the exhaust-gas system reduces the pollutants in the exhaust gas that are generated by an internal-combustion engine. The exhaust-gas system also helps to muffle exhaust-gas noise and to discharge the exhaust gas at a convenient point on the vehicle. Engine power should be reduced as little as possible during the process.

Components
An exhaust-gas system consists of the exhaust manifold, the components for exhaust-gas treatment and for sound absorption, and the connections between these components.

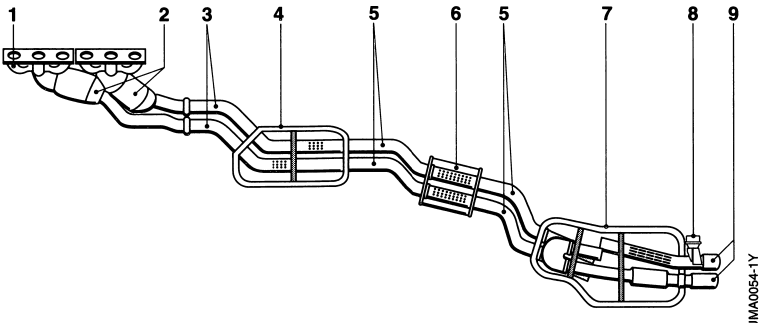
The design and the layout of these components are very different for passenger cars and commercial vehicles. In the case of passenger-car exhaust-gas systems the components are usually individually connected by pipes and routed under the vehicle chassis (Figure 1). Depending on engine displacement and the type of muffler used, a passenger-car exhaust-gas system weighs between 8 and 40 kg. The components are primarily made of high-alloy steels, as they are exposed to corrosion attacks from the inside by hot gas and condensate, and from the outside by moisture and splash and salt water.

Since the introduction of Euro IV, commercial vehicles have also been required to carry components for exhaust-gas treatment. These are, however, usually integrated into a large system and secured to the frame (see Commercial-vehicle exhaust-gas systems). The individual components are explained in the following using a passenger-car exhaust-gas system as the example

Exhaust-gas treatment
The components of the exhaust-gas treatment include
- the catalytic converter to break down the gaseous pollutants in the exhaust gas, and
- the particulate filter (or soot filter) to filter out the fine, solid particulates in the exhaust gas (especially in diesel engines).

Catalytic converters are installed in the exhaust-gas system as close as possible to the engine so that they can quickly reach their operating temperature and therefore be effective even at low operating temperatures (e.g. in urban driving). The decisive factor here is the light-off temperature of the catalytic converter, i.e. the temperature at which the catalytic converter begins to break down pollutants (approx. 250 °C for three-way convert-

Figure 1: Exhaust-gas system of a passenger car (example with three mufflers)
1 Exhaust manifold, 2 Near-engine catalytic converter, 3 Front pipe, 4 Front muffler, 5 Intermediate pipe, 6 Center muffler, 7 Rear muffler, 8 Exhaust-gas flap, 9 Tailpipe.



ers). Some catalytic-converter coatings, such as those for NO_x storage converters for example, are highly sensitive to temperature. These catalytic converters are therefore sooner installed in the cooler underbody area.

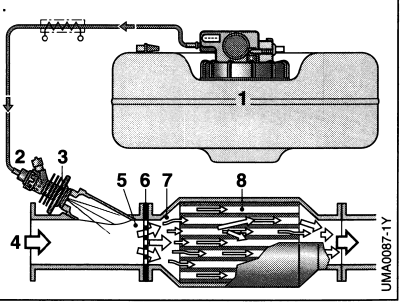
Diesel particulate filters are also installed in the front area of the exhaust-gas system to ensure that the soot particles they have retained are burnt off more effectively at the higher exhaust-gas temperatures.

Essentially three-way catalytic converters are used in spark-ignition engines. During engine operation with a stoichiometric air-fuel mixture ($\lambda = 1$) they convert the hydrocarbons (HC), nitrogen oxides (NO_x) and carbon monoxide (CO) present in the untreated exhaust gas by more than 99 %. An additional NO_x storage catalytic converter is only installed in lean-burn-operated gasoline direct-injection engines. This reduces the nitrogen oxides (NO_x) increasingly produced in this operating state.

Diesel engines require an oxidation-type catalytic converter to oxidize the hydrocarbons (HC) and carbon monoxide (CO) and a particulate filter to hold back the solid exhaust-gas components. For low NO_x emissions a further catalytic converter for NO_x reduction is required – an NO_x storage catalytic converter or an SCR catalytic converter (Selective Catalytic Reduction).

In order to ensure that the emissions are converted as much as possible in the catalytic converter or are well filtered in the particulate filter, it is essential to optimize the flow against these components in the exhaust-gas system. This is generally achieved by the shaping of the inflow funnel. Additional components such as swirl or mixing elements are required to optimize uniform distribution further. Especially SCR systems in which a reducing agent (Urea/water solution) is injected in liquid form into the exhaust-gas system require a mixer in order to ensure vaporization or atomization of the reducing agent to NH_3. The aim is to ensure uniform and gaseous distribution of NH_3 ahead of the catalytic converter (Figure 2).

Figure 2: Atomization and vaporization of reducing agent in SCR system
1 Reservoir for reducing agent,
2 Connection, injector for reducing agent,
3 Injector for reducing agent,
4 Gas flow, 5 Exhaust pipe, 6 Mixer,
7 Inflow funnel, 8 SCR catalytic converter.



Sound absorption

The primary cause of exhaust-gas noise is the gas pulsations of the internal-combustion engine, i.e. the gas vibrations which are generated by the combustion process and by the exhaust gas being forced out through the exhaust valves during each engine power cycle. This pulsation noise is in fact damped slightly by the catalytic converters and particulate filter. But this is not sufficient to keep the noise below the pass-by limit value prescribed by law (see Vehicle acoustics). For this reason, mufflers are installed in the middle or rear section of the exhaust-gas system.

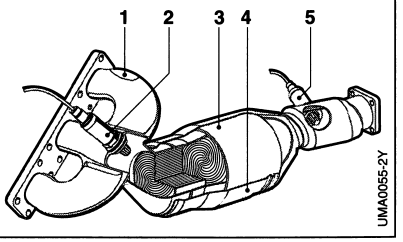
Depending on the number of cylinders and engine output, generally one, two or three mufflers are used in an exhaust-gas system. In V-engines, the left and right cylinder banks are often run separately, each being fitted with its own catalytic converters and mufflers.

The noise-emission limit for the complete vehicle is defined by legislation. The noise produced by the exhaust-gas system represents a substantial source of noise emission in a vehicle. This fact makes it necessary to devote particular attention and resources to the development of mufflers. Although the aim is to reduce noise in compliance with the legislation, they can also create the sound specific to the type of vehicle (sound design).

Exhaust manifold

An important component in the exhaust-gas system is the exhaust manifold (Figure 3). It routes the exhaust gas out of the cylinder outlet ports into the exhaust-gas system. The geometric design of the manifold (i.e. length and cross-section of the individual pipes) has an impact on the performance characteristics, the acoustic behavior of the exhaust-gas system, and the exhaust-gas temperature. In some cases, the exhaust manifold is air-gap-insulated (i.e. double-walled) in order to achieve high exhaust-gas temperatures quickly so as to enable the catalytic converter to work at an early stage during cold start. The exhaust manifold must consistently endure very high temperatures (up to 1,050 °C in gasoline engines). It is therefore produced from very high-quality material (high-alloy casting or high-alloy stainless steel).

Figure 3: Exhaust manifold with near-engine catalytic converter
1 Exhaust manifold, 2 λ sensor, 3 Metal monolith, 4 Insulating shell, 5 OBD sensor.



Catalytic converter

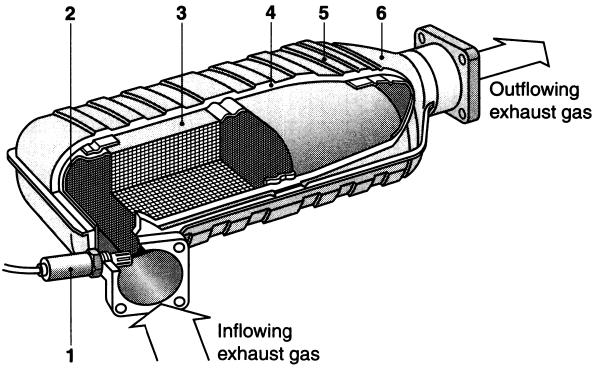
A catalytic converter is made up of an inflow funnel, an outflow funnel and a monolith (Figure 4). The monolith contains a large number of very fine, parallel channels covered with an active catalytic coating. The number of channels ranges from 400 to 1,200 cpsi (cells per square inch). The functional principle of the active catalytic-converter layer is described elsewhere (see Catalytic exhaust-gas treatment). For operational reasons, several differently coated monoliths are often used in one catalytic converter. Particular attention must be paid to the shape of the inflow funnel to ensure that the exhaust gas flows uniformly through the monolith. The external shape of the monolith depends on the installation space available in the vehicle, and may be triangular, oval or round.

The monolith can be made of metal or ceramic material.

Metal monolith

The metal monolith is made of finely corrugated, 0.05 mm thick metal foil, wound and brazed in a high-temperature process. Due to the very thin walls between the channels, the metal monolith offers an extremely low resistance to the exhaust gas. It is therefore frequently used in high-performance vehicles. The metal monolith can be welded directly to the funnels.

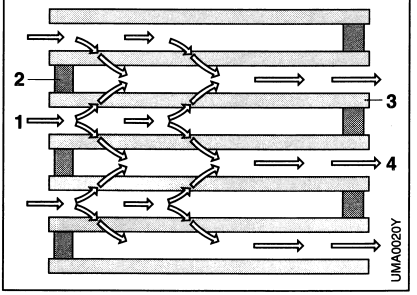
Figure 4: Catalytic converter with ceramic monolith
1 λ sensor,
2 Inflow funnel,
3 Ceramic monolith,
4 Mounting mat,
5 Metal housing,
6 Outflow funnel.



Ceramic monolith
The ceramic monolith is made of cordierite. Depending on the cell density, the wall thickness between the channels ranges from 0.05 mm (at 1,200 cpsi) to 0.16 mm (at 400 cpsi).

Ceramic monoliths have an extremely high stability to temperature and thermal shock. However, they cannot be installed directly in the metal housing and require a special mounting. This mounting is necessary in order to compensate for the difference between the thermal-expansion coefficients of steel and ceramics, and to protect the sensitive monolith against shocks. Extreme care and attention is required in the mounting and production process, particularly for thin-walled monoliths (< 0.08 mm). The monolith is mounted on a mat located between the metal housing and the ceramic monolith. The mounting mat is made up of ceramic fibers. It is extremely flexible to ensure that the pressure load exerted on the monoliths is at a minimum. It also serves as a heat insulator.

Figure 5: Ceramic particulate filter
1 Exhaust-gas inlet,
2 Ceramic plug,
3 Cell partition,
4 Exhaust-gas outlet.



Particulate filter

As with catalytic-converter monoliths, there are metallic and ceramic particulate-filter systems. The method of installing and mounting this filter in the metal housing is the same as the process used for the catalytic converter.

Design
In the same way as the ceramic monolith for the catalytic converter, the ceramic particulate filter is made up of a large number of parallel channels. However, these channels are alternately closed (Figure 5). Consequently, the exhaust gas is forced to flow through the porous walls of the honeycomb structure. The solid particles are deposited in the pores. Depending on the porosity of the ceramic body, the filtration efficiency of these filters can attain up to 97 %.

Regeneration
The soot deposits in the particulate filter induce a steady rise in flow resistance. For this reason, the particulate filter must be regenerated at certain intervals in two different processes. Details on these processes are described elsewhere (see Exhaust-gas treatment for diesel engines).

Passive process
In the passive process, the soot is burnt off by a catalytic reaction. For this purpose, an additive in the diesel fuel reduces the flammability of the soot particles to normal exhaust-gas temperatures.

Other passive regeneration options include catalytically coated particulate filters or the CRT process (see Continuous Regeneration Trap).

Active process
In the active process, external measures are implemented to heat the filter to the temperature necessary for burning off the soot. This rise in temperature can be achieved by a burner mounted upstream of the filter or by secondary injection initiated by the engine management and the use of a preliminary catalytic converter.

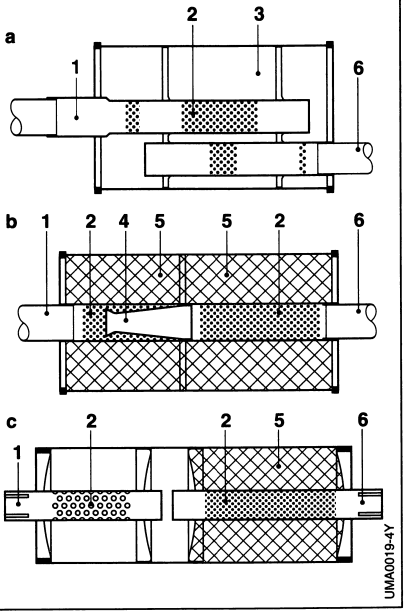
Mufflers

Mufflers (or silencers) are intended to smooth out exhaust-gas pulsations and make them as inaudible as possible. There are essentially two physical principles involved: reflection and absorption. Mufflers also differ according to these principles. However, they mostly comprise a combination of reflection and absorption.

In a reflection muffler (Figure 6a) the sound waves travel through the pipes and are reflected at the pipe outlet in the chambers. In an absorption muffler (Figure 6b) the sound waves spread over the perforation in the absorption material, where they are then muffled. Figure 6c shows a combination of reflection muffler and absorption muffler.

As mufflers and the exhaust-gas system pipes together form an oscillating system with its own natural resonance, the position of the mufflers is highly significant for the quality of sound-damping. The objective is to tune the exhaust-gas systems as low as possible, so that their natural frequencies do not excite bodywork resonances. To avoid structure-borne noise and to provide heat insulation for the vehicle underbody, mufflers often have double walls and an insulating layer.

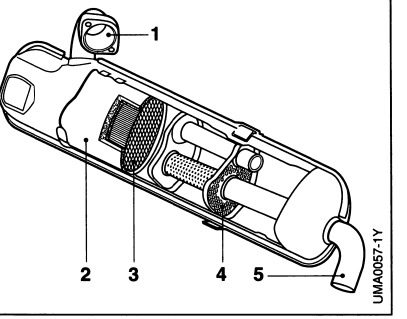
Figure 6: Muffler principles
a) Reflection muffler,
b) Absorption muffler,
c) Combination of reflection and absorption mufflers.
1 Inlet pipe,
2 Perforated pipe,
3 Chamber,
4 Venturi nozzle,
5 Absorption material,
6 Outlet pipe.



Reflection mufflers

Reflection (resonator-type) mufflers consist of chambers of varying lengths which are interconnected by pipes (Figure 6a and Figure 7). The pipes and partitions are perforated, thereby allowing exhaust gases to pass through. The differences in the cross-sections between the pipes and the chambers, the diversion of the exhaust gases, and the resonators formed by the connecting pipes and the chambers generate superpositions of the sound waves moving back and forth, causing the sound waves to be partially extinguished. In this way, effective muffling is achieved, particularly in the middle and low frequency ranges. The more such chambers are used, the more efficient is the muffler.

Figure 7: Muffler with integrated catalytic converter
1 Inlet pipe,
2 Mounting mat,
3 Ceramic monolith,
4 Reflection muffler,
5 Tailpipe.



Absorption mufflers

Absorption mufflers are designed with one chamber, through which a perforated pipe passes (Figure 6b). The chamber is filled with absorption material (mineral wool or glass fibers). The sound enters the absorption material through the perforated pipe and is converted into heat by friction.

The absorption material usually consists of long-fiber mineral wool or glass fibers with a bulk density of 100 to 150 g/*l*. The level of muffling depends on the bulk density, the sound-absorption grade of the material, and on the length and coating thickness of the chamber. Damping takes place across a very broad frequency band, but only begins at higher frequencies.

The shape of the perforations, and the fact that the pipe passes through the wool, ensures that the material is not blown out by exhaust-gas pulses. Sometimes the mineral wool is protected by a layer of stainless-steel wool around the perforated pipe.

Because the exhaust gas in absorption mufflers essentially passes through a straight pipe, the pressure loss here is very low compared with the pressure loss at the differences in the cross-sections in reflection mufflers.

Muffler design

Depending on the space available under the vehicle, mufflers are produced either as spiral-wound casing or from half-shells.

To produce the jacket for a spiral-wound muffler, one or several metal sheet blanks are shaped over a round mandril and joined together either by the longitudinal folds or by laser welding. The completely assembled and welded core is then installed in the jacket casing. It consists of internal tubes, baffles, and intermediate layers. The outer layers are then connected to the jacket in a folding or laser-welding process.

It is often not possible to effectively accommodate a spiral-wound muffler in view of the complicated space conditions in the floor assembly. In such cases, a shell-type muffler made of deep-drawn half-shells is used as it can assume virtually any required shape.

The total volume of the mufflers in a passenger-car exhaust-gas system corresponds to approximately eight to twelve times the engine displacement.

Connecting elements

Pipes are used to connect the catalytic converters and mufflers together. Arrangements where the catalytic converter and muffler are integrated in a single housing may also be used on very small engines and vehicles (Figure 7).

The pipes, the catalytic converter and muffler are connected to form an integrated system by means of plug-in connections and flanges. Many original-equipment systems are fully welded for faster mounting.

The entire exhaust-gas system is connected to the vehicle underbody via elastic mounting elements (Figure 8). The fixing points must be carefully selected, otherwise vibration may be transmitted to the bodywork and generate noise in the passenger compartment. The wrong fixing points may also create strength and therefore durability problems. In some cases, these problems are counteracted by the use of vibration absorbers. These components oscillate at the critical frequency in precisely the opposite direction of the exhaust-gas system, thereby absorbing system vibration.

The exhaust-gas system noise at the exhaust-emission point (tailpipe), as well as sound radiation from the mufflers, can also cause bodywork resonance. Depending on the intensity of the engine vibrations, decoupling elements (Figure 9) are used to isolate the exhaust-gas system from the engine block and to relieve the stress load on the exhaust-gas system. A decoupling element comprises the liner, i.e. an inner pipe with pipe segments that move in each other. The corrugated pipe, also known as the boot, is mounted on top. The liner serves to protect the boot and limit linear expansion. The corrugated pipe exercises the decoupling/isolating effect by way of its soft structure. The corrugated pipe is protected against external influences by a wire-fabric jacket.

Finally, the mounting arrangement of an exhaust-gas system is tuned so that it is rigid enough to withstand vibrations reliably on the one hand, and it exhibits sufficient flexibility and damping properties to reduce the transfer of forces to the bodywork effectively on the other.

Figure 8: Mounting element
1 Rubber mount,
2 Metal bracket,
3 Muffler shell.

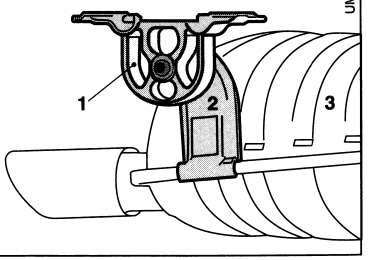
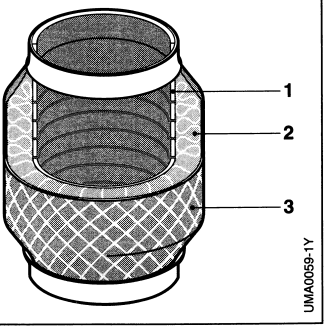


Figure 9: Decoupling element
1 Liner,
2 Corrugated pipe (boot),
3 Wire-fabric jacket.



Acoustic tuning devices

A number of different components can be used to eliminate disturbing frequencies in the noise emitted from the tailpipe. These components can be used specifically to damp individual frequency ranges very effectively.

Helmholtz resonator
The Helmholtz resonator consists of a pipe arranged along the side of the exhaust-gas train and a defined volume connected to it (Figure 10). The gas volume acts as a spring, while the gas in the pipe section acts as a mass. At its resonant frequency, this spring-mass system provides a very high degree of sound absorption but in a narrow frequency band. The resonant frequency f is dependent on the size of the volume V as well as the length L and the cross-sectional area A of the pipe:

$$f = \frac{c}{2\pi}\sqrt{\frac{A}{L \cdot V}}$$

The value c is the speed of sound.

λ/4 resonators
λ/4 resonators consist of a pipe, sealed at the end, branching off from the exhaust-gas system. The resonant frequency f of these resonators is derived from the length L of the pipe branch. It is expressed as:

$$f = \frac{c}{4L}.$$

These resonators also feature a very narrowband damping range about their resonant frequency.

Exhaust-gas flaps
Exhaust-gas flaps are most commonly found in rear mufflers. Depending on the engine speed or exhaust-gas throughput, they close off a bypass pipe in the muffler or a second tailpipe (Figure 11). As a result, exhaust-gas noise can be substantially damped at lower engine speeds without the need to trade off power losses at high engine speeds.

Exhaust-gas flaps can be either self-controlling based on pressure and flow, or they can be controlled externally. An interface to the engine management system must be provided for externally controlled flaps. This makes them more complex than self-controlled flaps. However, their application range is also more flexible.

Figure 10: Helmholtz resonator
1 Helmholtz volume,
2 Gas-carrying exhaust pipe,
3 Helmholtz pipe,
4 Muffler shell.

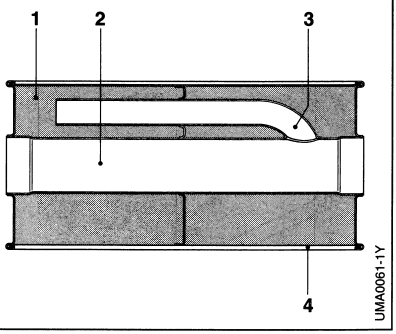
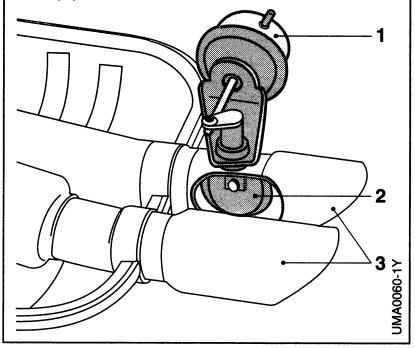


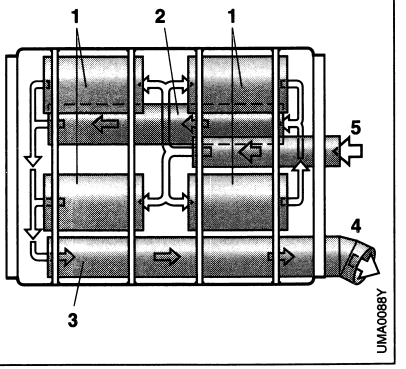
Figure 11: Exhaust-gas flap, controlled by vacuum
1 Vacuum unit,
2 Butterfly valve,
3 Tailpipes.



Commercial-vehicle exhaust-gas systems

In commercial-vehicle exhaust-gas systems most of the previously described components are integrated in a housing which is mounted to the vehicle frame. The number of catalytic converters and particulate filters depends on which emissions-control legislation the exhaust-gas system is designed for.

Figure 12: Commercial-vehicle exhaust-gas system for Euro IV
1 SCR catalytic converters, 2 Inlet pipe,
3 Outlet pipe, 4 Tailpipe, 5 Exhaust-gas inlet.

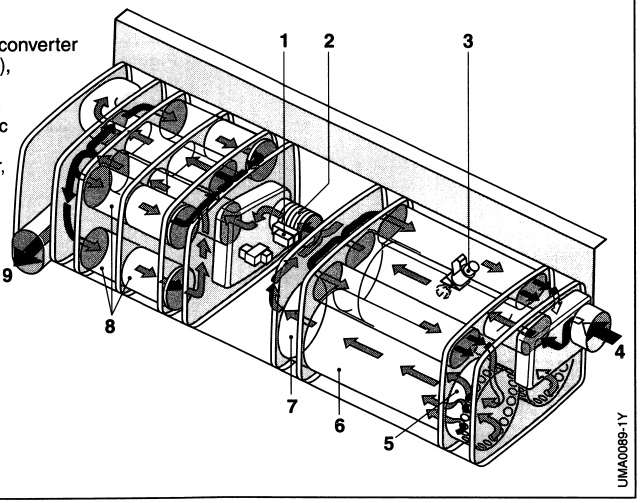


Exhaust-gas systems for Euro IV and Euro V

Generally speaking, a particulate filter is not required for Euro IV and Euro V. Only oxidation and SCR catalytic converters are used (see Exhaust-gas treatment for diesel engines). Alternatively, the diesel engine can also be tuned in such a way that the untreated NO_x emissions comply with the limit values for Euro IV and Euro V. But then a particulate filter is required in the exhaust-gas system.

Figure 12 shows a Euro IV exhaust-gas system with SCR catalytic converters. In contrast to passenger-car exhaust-gas systems, several catalytic converters are frequently arranged in parallel in order to accommodate the required catalytic surface in the space available. Bypass pipes and holes in the bases serve to route the gases and muffle the sound. Depending on the engine size, these systems attain a volume of 150 to 200 *l* and a weight of 150 kg.

Figure 13: Commercial-vehicle exhaust-gas system for Euro VI and EPA 10
1 Urea injector,
2 Decoupling element,
3 HC injector,
4 Inlet pipe,
5 Pre-oxidation catalytic converter (pre catalytic converter),
6 Diesel particulate filter,
7 Post-oxidation catalytic converter (post catalytic converter),
8 SCR catalytic converter,
9 Outlet pipe.



Exhaust-gas systems for EuroVI and EPA 10

Exhaust-gas systems for the latest legislation (Euro IV in Europe and EPA 10 in the USA) require all the components, i.e. oxidation catalytic converters, particulate filters and SCR catalytic converters (Figure 13). They are thus even bigger in terms of volume and weight.

Two concepts are currently used. Either all the components are accommodated in a single housing or the SCR catalytic converters and the particulate filters are divided between two housings. The following components are additionally required to ensure correct functioning of exhaust-gas treatment: The SCR function requires a urea metering system (see SCR system), whose nozzle (injector) is located at a suitable location in the exhaust-gas system. Furthermore, to ensure reliable regeneration of the particulate filter, an HC metering unit is frequently needed to inject fuel (see HCI system). In both cases a position must be chosen which ensures that the liquid urea solution and the fuel are adequately vaporized and distributed. If necessary, mixers are used here to prepare the liquids better (see section on Exhaust-gas treatment).

A variety of sensors must also be accommodated in the housings. In addition to pressure sensors for monitoring the filter load, temperature sensors and NO_x sensors for monitoring NO_x conversion are required. The position of the sensors in the exhaust-gas system must, depending on the system design, be optimized to guarantee an adequate signal quality in all operating states.

References

[1] C. Hagelüken: Autoabgaskatalysatoren – Grundlagen, Herstellung, Entwicklung, Recycling, Ökologie. 3rd Edition, Expert-Verlag, 2016.

Management for spark-ignition engines

Description of the engine management system

The engine management system ensures that the driver command is implemented. The driver can request acceleration, deceleration or driving at a constant speed – the engine management system ensures that the spark-ignition engine's drive output required to achieve this is set. The system controls all engine functions in such a way that the engine delivers the required level of torque, but fuel consumption and exhaust emissions are kept low.

The power output from a spark-ignition (SI) engine is determined by the available clutch torque and the engine speed. The clutch torque is the torque developed by the combustion process less friction torque (friction losses in the engine), pumping losses, and the torque needed to drive the auxiliary equipment (Figure 1). The drive torque is available at the wheels. This is developed from the clutch torque reduced by the losses in the clutch and in the transmission. This resulting torque is offset by the tractive resistances such as rolling resistance and aerodynamic drag. Depending on the driver command, a state of equilibrium or imbalance exists between these resistances and the drive torque. When the resistances and the drive torque are equal, a constant driving speed is adopted, otherwise the vehicle is subject to acceleration or deceleration.

Combustion torque is generated in the engine's power cycle and is mainly determined by the following variables:

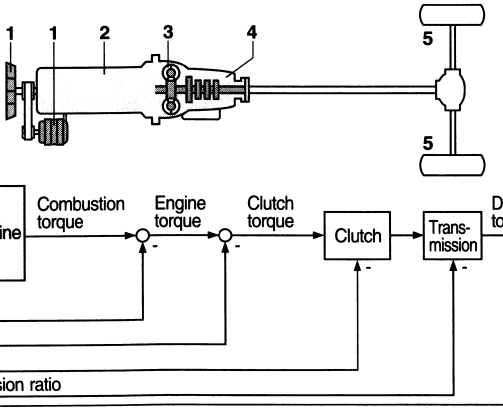
- The air mass which is available for combustion when the intake valves close,
- The fuel mass available in the cylinder,
- The point at which the combustion takes place.

Furthermore, there are also minor influences, for example, due to mixture composition (amount of residual exhaust gas) or combustion processes.

The primary function of engine management is to coordinate the various subsystems (air, fuel and ignition systems) in order to adjust the torque requested by the engine and at the same time meet the exacting standards placed on exhaust-gas emissions, fuel consumption, power output, comfort, and safety. The engine-management system also performs diagnoses on the subsystems.

Figure 1: Drivetrain torques

1 Auxiliary systems (alternator, A/C compressor, etc.),
2 Engine,
3 Clutch,
4 Transmission,
5 Drive wheels.



System overview

Electrical system overview

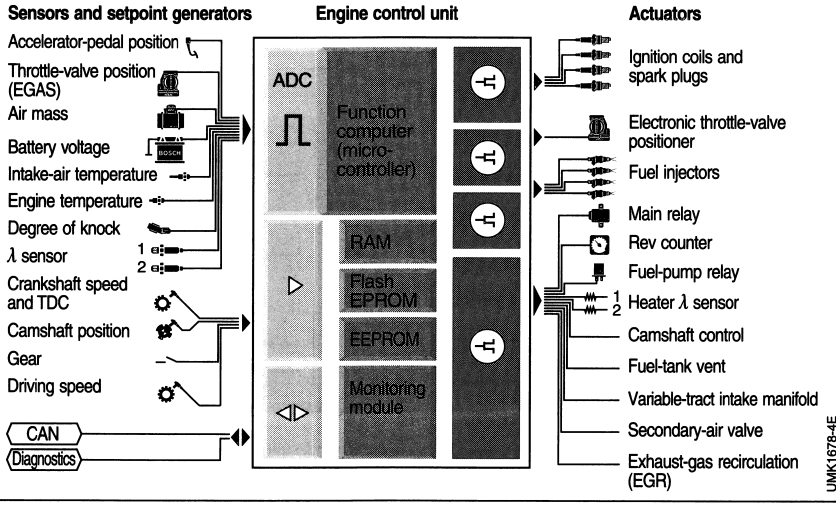
Motronic is the name given to Bosch systems for the open- and closed-loop control of spark-ignition (SI) engines. The Motronic system (Figure 2) comprises all the sensors for detecting ongoing operating data from the engine and vehicle, and all the actuators that perform the required adjustments to the SI engine. The control unit uses data from the sensors to scan the status of the vehicle and engine at very short intervals (every few milliseconds to meet the system's real-time requirements). Input circuits suppress sensor-signal interference and convert the signals to a single uniform voltage scale. An analog/digital converter then transforms the conditioned analog signals into digital values. Other signals are received via digital interfaces (e.g. CAN bus, FlexRay) or via pulse-width-modulated (PWM) interfaces.

The nerve center of the engine control unit is a microcontroller with a program memory (e.g. flash EPROM), which stores all the process control algorithms – i.e. arithmetical processes performed according to a specific pattern – and data (parameters, characteristics, program maps). The input variables derived from the sensor signals influence the calculations in the algorithms, and thus the triggering signals for the actuators. From these input signals, the microcontroller detects the sort of vehicle response the driver wants and calculates, for example, the required torque, the resulting cylinder charge with the associated injected fuel quantity, the correct ignition timing, and the triggering signals for the actuators (e.g. the evaporative-emissions control system, the exhaust-gas turbocharger, and the secondary-air system).

The low-level signal data from the microcontroller outputs are adapted by the driver stages to the levels required by the various actuators.

Another important function of Motronic is the monitoring of the operational ability of the entire system, using On-Board Diagnostics (OBD). Legal requirements (diagnosis legislation) have placed extra demands on Motronic, with the result that about half of Motronic system capacity (in terms of computing power and memory requirements) is dedicated to diagnostics-related tasks.

Figure 2: Components used for electronic control of a spark-ignition engine



Functional system overview

An engine management system has many other secondary functions in addition to its primary functions of cylinder-charge control, fuel supply, mixture formation, and ignition. The entire system is broken down into subsystems so that they can be described more clearly and in greater detail. The full description is shown in the system structure (Figure 3).

Torque Demand (TD) subsystem

The driver issues a concrete driver command through the accelerator-pedal position. The accelerator-pedal position is converted into a setpoint value for the drive torque.

As well as a direct torque input, the driver can also input an indirect driver command via the cruise-control system. A setpoint drive torque is calculated, depending on the current driving condition.

When the accelerator pedal is not pressed, that amount of engine torque which is needed to maintain idle speed is calculated.

Drivability filters, the surge-damping function, the electrical system (starter, alternator, battery) and other electrical consumers such as the air-conditioning system issue further torque demands.

Torque Structure (TS) subsystem

The varied torque demands from the Torque Demand subsystem, demands from the transmission system, driving-dynamics demands, and other engine-specific torque demands (e.g. catalytic-converter heating) are coordinated in the Torque Structure subsystem. The result is the generation of a torque demand made of the internal-combustion engine.

Setpoint values for charge, injection and ignition are generated from the resulting torque demand of the engine.

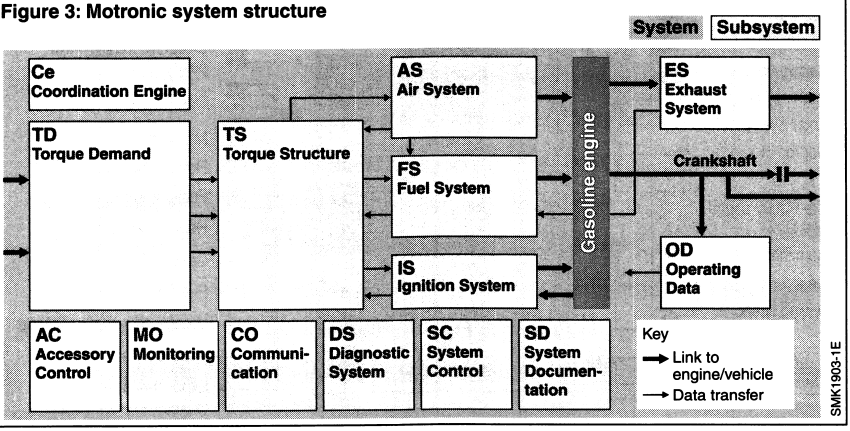
The charge is input as a relative air mass. The relative air mass (standardization for all engine-performance classes) is the ratio between the real air mass in the cylinder and the maximum possible air mass in the cylinder at the current engine speed.

The setpoint value for ignition is described by an ignition angle.

Torque reductions can be represented with injection blank-outs (e.g. traction control system demand). The number of blank-outs is determined for this purpose.

In systems with gasoline direct injection, it is possible to set lean-burn operating modes (e.g. stratified fresh charge in the combustion chamber). In these operating modes, it is also possible to set the engine torque by inputting a λ setpoint value.

Figure 3: Motronic system structure



Physical models calculate the actual engine torque at the clutch from various sensor signals. The actual torque is used to monitor the Motronic system and is required by other systems such as transmission control.

Air System (AS) subsystem
The input of the relative setpoint air mass from the Torque Structure subsystem is converted into concrete quantities for the actuators which are used for cylinder-charge control.

The throttle valve is the main actuator for charge. Models are used to calculate the throttle-valve opening angle from the setpoint air mass and from this opening angle, a pulse-width-modulated activation of the actuator.

There are systems in which the main setting path is represented by the activation of intake and exhaust valves. The throttle valve usually remains constantly open in such systems. Only in special situations (e.g. in limp-home mode) is the throttle valve used in these systems as the setting path for charge.

In turbocharged engines, wastegate activation for the exhaust-gas turbocharger or control of mechanical superchargers are also taken into consideration.

The camshaft adjustment systems and exhaust-gas recirculation valves are further setting devices.

In addition, the current actual charge of the internal-combustion engine is determined. Sensor signals such as pressure and temperature in the intake manifold are used as basic variables for this purpose.

Fuel System (FS) subsystem
The function of the fuel system is to supply the fuel rail with fuel from the fuel tank in the requested quantity and at the prespecified pressure.

By using the current actual charge, the fuel pressure in the rail and the pressure in the intake manifold, the opening duration of the fuel injectors is calculated from the λ setpoint value.

The fuel injectors are activated synchronously with the crank angle in order to optimize the air/fuel mixture.

Longer-term adaptations of the actual value for λ ensure greater pilot-control accuracy of fuel metering.

Ignition System (IS) subsystem
The resulting ignition angle is calculated from the setpoint input for ignition, from the engine operating conditions and on the basis of interventions (e.g. knock control), and an ignition spark is generated at the spark plug at the desired moment of ignition.

The ignition angle is set in such a way that the engine runs with optimal consumption. The subsystem only deviates from this in a few special situations (e.g. catalytic-converter heating or fast torque reduction when changing gear).

The knock-control system permanently monitors combustion in all the cylinders. It ensures that the engine runs with optimal consumption close to the knock limit. At the same time, damage caused by combustion knock is safely avoided. Failures in the knock-detection path are subject to continuous monitoring so that ignition can be effected in the event of a fault at sufficient distance from the knock limit.

Exhaust System (ES) subsystem
The open- and closed-loop control interventions for optimum operation of the three-way catalytic converter are calculated in this subsystem. The combustion mixture must be regulated within narrow limits around the stoichiometric mixture ratio.

The operational capability of the catalytic converter is also monitored. The signals from exhaust-gas sensors (e.g. λ oxygen sensor) serve as the basis for this monitoring.

Component protection functions ensure that the exhaust system is not subject to thermal overloads. The actual temperatures in the exhaust system required for this purpose are usually modeled.

In lean-burn operating modes with stratified fresh charge (in the case of gasoline direct injection), the combustion mixture is also controlled to ensure optimum operation of the NO_x storage catalytic converter.

Coordination Engine (CE) subsystem
In the case of gasoline direct injection, the operating modes (e.g. operation with homogeneous or stratified mixture distribution in the combustion chamber) are coordinated and switched over. In order to determine the required operating mode, the demands of various functionalities must be coordinated on the basis of defined priorities.

Operating Data (OD) subsystem
The Operating Data subsystem evaluates engine operating-condition variables (e.g. engine speed, temperatures), ensures digital conditioning and plausibilization, and makes the result available to the other subsystems.

Adaptation of tolerances in speed sensing helps to control injection and ignition more accurately.

Engine misfires – as a precondition for catalytic-converter protective functions – are detected.

Accessory Control (AC) subsystem
Additional functionalities such as A/C-compressor control, fan control or engine-temperature regulation are frequently integrated into the engine management system. They are coordinated in the Accessory Control subsystem.

Communication (CO) subsystem
The vehicle network contains many other systems (e.g. transmission control, Electronic Stability Program), as well as the Motronic system. The systems exchange information via standardized interfaces (e.g. CAN communication).

In addition, workshop testers can read out signals from the engine management system and perform defined actuator adjustments (actuator diagnosis).

Diagnostic System (DS) subsystem
The operational capability of the Motronic system is subject to continuous monitoring by diagnostic functions. These include both electrical and plausibility checks though comparisons of sensor signals with models. Faults are stored and managed (e.g. faults are assigned "time stamps"). The faults can be called up at a later stage on workshop testers.

Some diagnostic functions are only operational under certain boundary conditions (e.g. in specific temperature or load ranges). There are also diagnostic function which must occur in a prespecified sequence. Coordination of this sequence control is performed by the Diagnostic System.

Monitoring (MO) subsystem
"Drive by wire" systems are relevant to monitoring. The core function is torque comparison. This compares the permissible torque calculated on the basis of driver command with the actual torque calculated from the engine data.

On further levels, the computer core and its peripherals are monitored.

System Control (SC) subsystem
The run-up of the Motronic system is regulated. Computing frameworks must be provided before individual functionalities are calculated. Different computing frameworks (e.g. frameworks with angle or time synchronism) are required for the resource optimization of the computing time.

Defined functionalities (e.g. function diagnosis of output stages) occur before the engine starts. Sequence control also handles resets and ECU run-on.

System Document (SD) subsystem
In addition to the extensive open- and closed-loop control functions of the Motronic system, numerous documents are required in order to describe a concrete project in detail. These documents include system descriptions of the ECU hardware and software, the wiring harness, the engine data, components, and connection-pin assignments.

Motronic versions for passenger cars

History

Originally, Motronic's function essentially combined electronic fuel injection and electronic ignition in a single control unit. Gradually, more and more functions were added in response to legislative requirements to reduce exhaust-gas emissions, reduce fuel consumption, and meet increased demands in terms of performance, driving comfort, and driving safety. Examples of these additional functions are:

- Idle-speed control,
- λ control,
- Control of the evaporative-emissions control system,
- Exhaust-gas recirculation to reduce NO_x emissions and fuel consumption,
- Control of the secondary-air system to reduce HC emissions in the starting and warm-up phases,
- Control of the exhaust-gas turbocharger and variable-tract intake manifold to increase engine performance,
- Camshaft control to reduce exhaust-gas emissions and fuel consumption and to increase performance,
- Component protection (e.g. knock control, engine-speed limitation, exhaust-gas temperature control).

The basis for the first Motronic, which was introduced from 1979 onwards in production vehicles, was the L-Jetronic fuel-injection system with air-flow sensing and electronic ignition with rotating distributor system for a manifold-injection engine with individual-cylinder injection and a conventional, mechanical throttle valve. This M-Motronic was further developed like the injection and ignition systems further used in parallel for this purpose. For load sensing the air-flow sensor was replaced by the air-mass meter, which enables the engine load to be sensed more accurately. Electronic ignition with rotating ignition distributor system was replaced by fully electronic ignition with static distributor system, in which a separate ignition coil generates the high voltage for each cylinder.

In addition to systems with electronic multipoint fuel injection, the following simpler and, at that time, cost-effective systems were also developed to allow Motronic to be used in mid-range and compact vehicles:

- KE-Motronic, based on KE-Jetronic continuous gasoline injection,
- Mono-Motronic, based on Mono-Jetronic intermittent single point injection.

State of the art

Now, only multipoint fuel-injection systems and fully electronic ignition are used for new vehicles. The following Motronic versions are used:

- ME-Motronic with electronic throttle control (ETC) for controlling injection, ignition, and fresh-air charge on intake-manifold injection systems (Figure 4).
- DI-Motronic (Direct Injection), with additional open- and closed-loop control functions for the high-pressure fuel circuit on gasoline direct-injection systems and to implement the various operating modes of this engine type (Figure 5 for homogeneous operation, Figure 6 for lean-burn concept).
- Bifuel-Motronic for controlling components for running the spark-ignition engine on natural gas or gasoline (see Engines fueled by natural gas).

Figure 4: ME-Motronic system diagram for manifold injection

1 Carbon canister, 2 Hot-film air-mass meter with integrated temperature sensor, 3 Throttle device (ETC), 4 Canister-purge valve, 5 Intake-manifold pressure sensor, 6 Fuel rail, 7 Fuel injector, 8 Actuators for variable camshaft control, 9 Ignition coil with attached spark plug, 10 Camshaft phase sensor, 11 λ sensor upstr. of close-coupled cat. conv. (two-step or broad-band sensor), 12 Engine ECU, 13 Exhaust-gas recirculation valve, 14 Speed sensor, 15 Knock sensor, 16 Engine-temperature sensor, 17 Close-coupled catalytic converter (three-way catalytic converter), 18 λ sensor downstr. of close-coupled cat. conv. (two-step sensor), 19 CAN interface, 20 Fault lamp, 21 Diagnosis interface, 22 Interface to immobilizer ECU, 23 Accelerator-pedal module with pedal-travel sensor, 24 Fuel tank, 25 In-tank unit with electric fuel pump, fuel filter and fuel pressure regulator, 26 Underfloor catalytic converter (three-way catalytic converter).

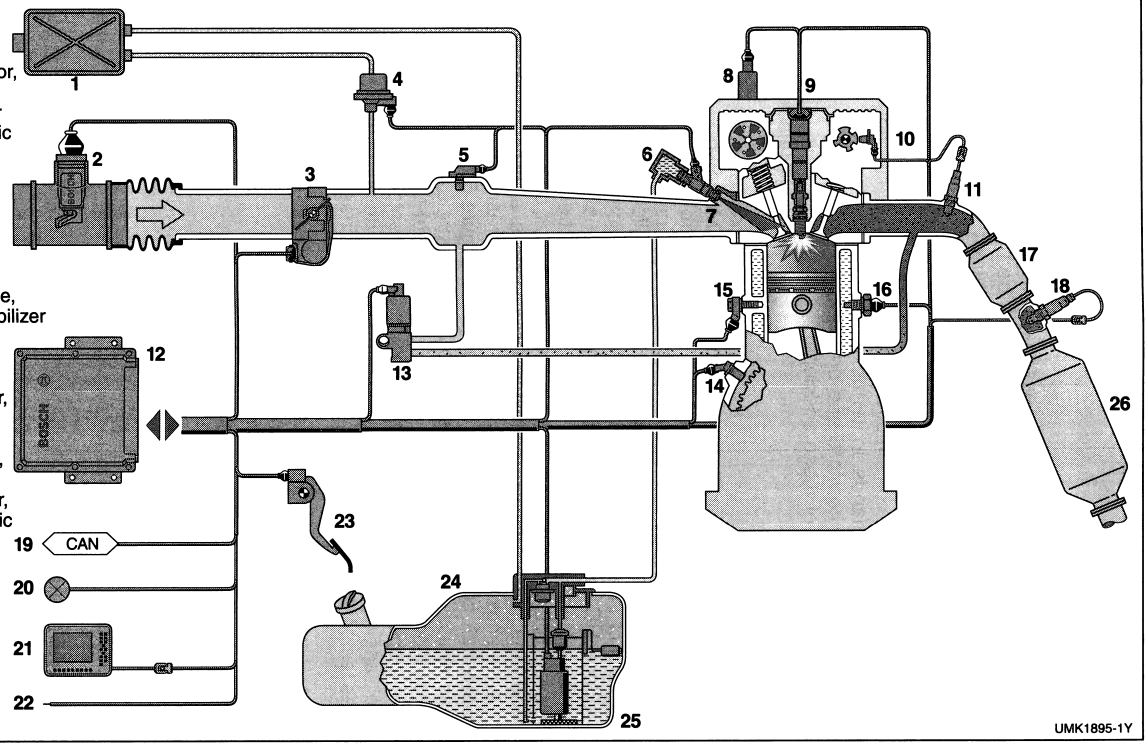


Figure 5: DI-Motronic system diagram for gasoline direct injection (homogeneous operation) with exhaust-gas turbocharging

1 Carbon canister, 2 Canister-purge valve, 3 Non-return valves, 4 Blow-off valve, 5 Hot-film air-mass meter with integrated temperature sensor, 6 Charge-air pressure sensor (optionally in combination with charge-air temperature sensor), 7 Intercooler, 8 High-pressure pump with integrated fuel-supply control valve, 9 Actuators for variable camshaft control, 10 Ignition coil with attached spark plug, 11 Vacuum pump for wastegate activation, 12 Camshaft phase sensor, 13 Ambient-pressure sensor, 14 Throttle device (ETC), 15 Swirl control valve, 16 Fuel-pressure sensor, 17 Fuel rail (high-pressure rail), 18 High-pressure fuel injector, 19 Engine-temperature sensor, 20 Exhaust-gas temperature sensor, 21 Boost-pressure control valve, 22 Wastegate (bypass valve), 23 Exhaust-gas turbocharger, 24 Engine ECU, 25 Speed sensor, 26 Knock sensor, 27 λ sensor upstr. of close-coupled cat. conv. (two-step or broad-band sensor), 28 CAN interface, 29 Fault lamp, 30 Diagnosis interface, 31 Interface to immobilizer ECU, 32 Accelerator-pedal module with pedal-travel sensor, 33 Fuel tank, 34 In-tank unit with electric fuel pump, fuel filter and fuel pressure regulator, 35 Close-coupled catalytic converter (three-way catalytic converter), 36 λ sensor downstr. of close-coupled cat. conv. (two-point sensor), 37 Underfloor catalytic converter (three-way catalytic converter).

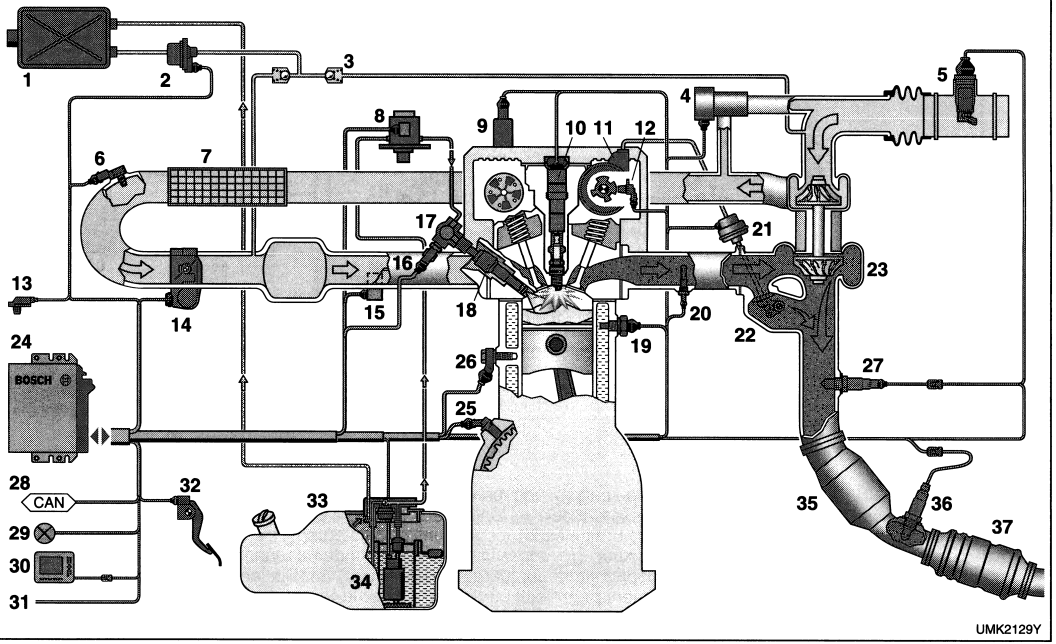
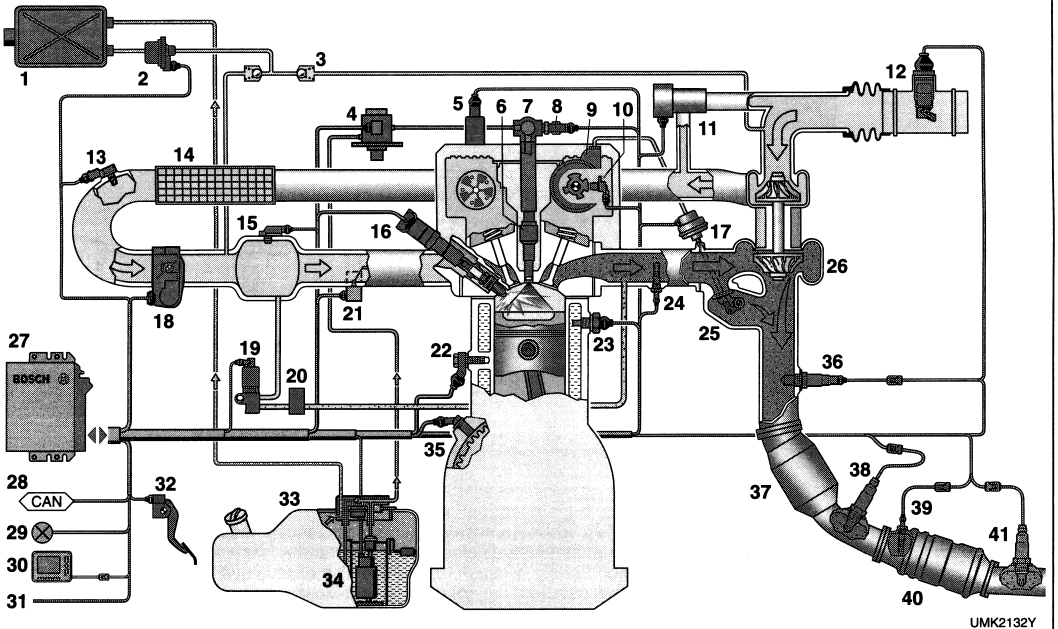


Figure 6: DI-Motronic system diagram for gasoline direct injection (lean-burn concept) with exhaust-gas turbocharging

1 Carbon canister, 2 Canister-purge valve, 3 Non-return valves, 4 High-pressure pump with integrated fuel-supply control valve, 5 Actuators for variable camshaft control, 6 High-pressure fuel injector (central arrangement for spray-guided combustion process), 7 Fuel rail (high-pressure rail), 8 Fuel-pressure sensor, 9 Vacuum pump for wastegate activation, 10 Camshaft phase sensor, 11 Blow-off valve, 12 Hot-film air-mass meter with integrated temperature sensor, 13 Charge-air pressure sensor (optionally in combination with charge-air temperature sensor), 14 Intercooler, 15 Intake-manifold pressure sensor, 16 Ignition coil with attached spark plug, 17 Boost-pressure control valve, 18 Throttle device (ETC), 19 Exhaust-gas recirculation valve, 20 Pressure sensor, 21 Swirl control valve, 22 Knock sensor, 23 Engine-temperature sensor, 24 Exhaust-gas temperature sensor, 25 Wastegate (bypass valve), 26 Exhaust-gas turbocharger, 27 Engine ECU, 28 CAN interface, 29 Fault lamp, 30 Diagnosis interface, 31 Interface to immobilizer ECU, 32 Accelerator-pedal module with pedal-travel sensor, 33 Fuel tank, 34 In-tank unit with electric fuel pump, fuel filter and fuel pressure regulator, 35 Speed sensor, 36 λ sensor upstr. of close-coupled cat. conv. (broad-band sensor), 37 Close-coupled catalytic converter (three-way catalytic converter or combination of three-way and NO_x storage catalytic converter), 38 λ sensor downstr. of close-coupled cat. conv. (two-step sensor), 39 Exhaust-gas temperature sensor, 40 Underfloor catalytic converter (NO_x storage catalytic converter), 41 λ sensor (two-step sensor) or NO_x sensor.



Motronic versions for motorcycles

The huge diversity of motorcycles available on the market internationally is reflected in the wide range of requirements imposed on engine management. Engine management has not by a long shot arrived on the scene in all segments and markets of the motorcycle sector. The Asian market in particular – India and China – is dominated by single-cylinder vehicles with low engine displacement under 150 cm^3 which are still entirely reliant on carburetor technology. Today, the persistent widespread use of air-cooled engines (see Engine cooling) is a particular characteristic of the motorcycle market.

Because of stricter, applicable legislation for example in Europe (see Emission-control and diagnosis legislation) motorcycles under 150 cm^3 already have electronic engine-management systems.

Motronic for multiple-cylinder engines
Adapted engine-management systems from the passenger-car sector are used for large-capacity multiple-cylinder engines in motorcycles.

Motronic for single-cylinder engines
For single-cylinder engines there are standalone solutions for engine management. A distinction is made between the following systems:
- MSE-M Motronic (SE, Small Engines) for controlling ignition and fuel injection on intake-manifold injection systems with mechanical throttle valves.
- MSE-E Motronic for controlling ignition, fuel injection and fresh-air charge on intake-manifold injection systems with electronic throttle valves (ETC).

Two particular features of these single-cylinder engines for motorcycles are the high rotational non-uniformities of the crankshaft (Figure 7) and the high intake manifold-pressure dynamics over a working cycle (see Cylinder charge). This behavior characteristic of single-cylinder engines can be used to implement the assignment of the engine position to the compression cycle or to the exhaust cycle without a camshaft speed sensor.

System simplifications in the small-engine segment play a key role in equipping the huge number of motorcycles in Asia with efficient engine management.

Requirements
A further requirement that sets apart engine management for motorcycles – both single- and multiple-cylinder motorcycles – from that for passenger cars is the higher rpm range to be covered between idle and maximum speeds. This imposes high demands on the computing capacity of Motronic.

The increasing requirements with regard to exhaust-gas quality and requirements from on-board diagnosis (see Emission-control and diagnosis legislation) will result in system-engineering changes in the future. Thus, a second, heated λ sensor downstream of the catalytic converter is used for example to diagnose catalytic-converter aging. The catalytic converter itself must be adapted in terms of its dimensions to the more stringent limits. The catalytic-converter heating strategy performed by the engine management makes an important contribution to heating the catalytic converter quickly.

Figure 7: Comparison of rotational non-uniformity for cars and motorcycles.
1 Three-cylinder engine for cars,
2 Single-cylinder engine for motorcycles.

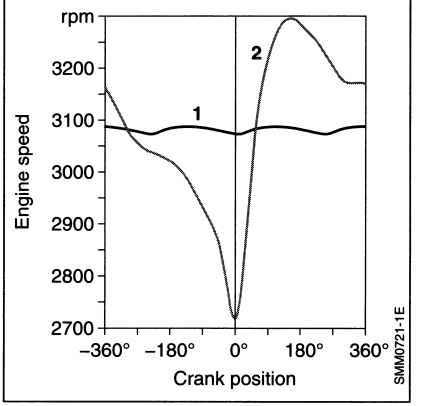


Figure 8: MSE-M Motronic with mechanical throttle valve for motorcycles
(System diagram with air-cooled engine).
1 Carbon canister, 2 Tank vent valve,
3 In-tank unit with electric fuel pump, fuel filter and fuel pressure regulator,
4 Fuel injector, 5 Ignition coil, 6 Spark plug, 7 rpm sensor,
8 Engine-temperature sensor, 9 λ sensor upstream of catalyst,
10 Intake-manifold pressure and temperature sensor, 11 Engine control unit, 12 CAN interface,
13 Fault lamp, 14 Diagnosis interface, 15 Mechanical throttle device,
16 Throttle grip with Bowden cable, 17 Throttle-valve sensor, 18 Controlled air bypass.

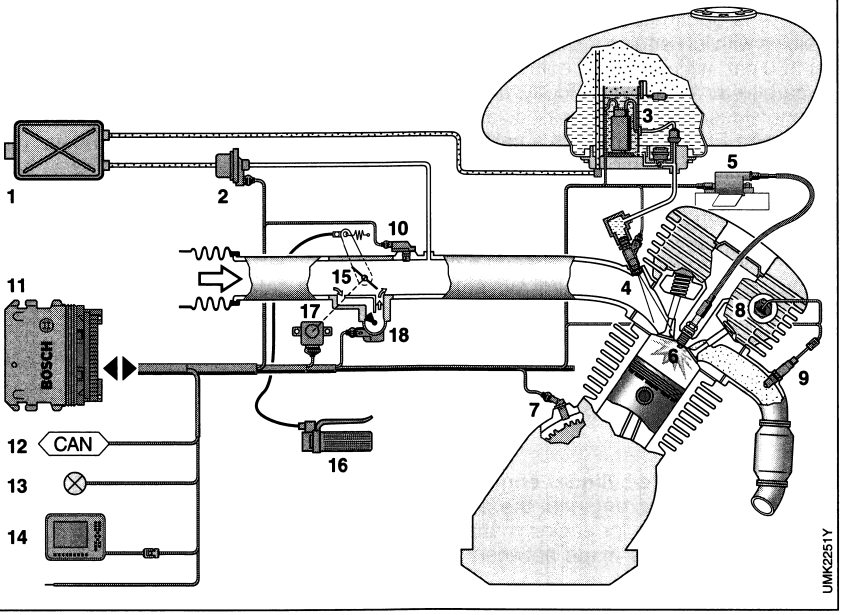
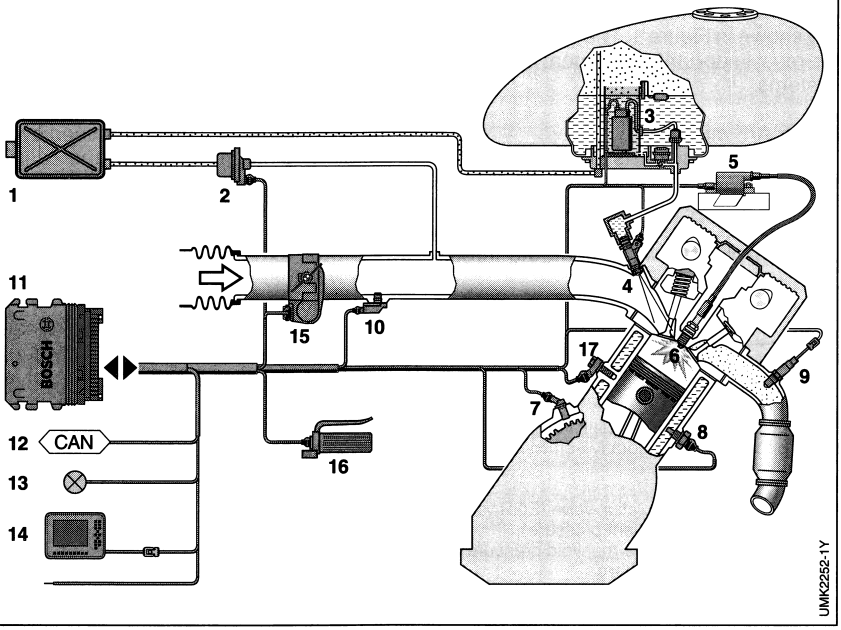


Figure 9: MSE-E Motronic with electronic throttle valve (ETC) for motorcycles
(System diagram with water-cooled engine).
1 Carbon canister, 2 Tank vent valve,
3 In-tank unit with electric fuel pump, fuel filter and fuel pressure regulator,
4 Fuel injector, 5 Ignition coil, 6 Spark plug, 7 rpm sensor,
8 Engine-temperature sensor, 9 λ sensor upstream of catalyst,
10 Intake-manifold pressure and temperature sensor, 11 Engine control unit, 12 CAN interface,
13 Fault lamp, 14 Diagnosis interface, 15 Electronic throttle device (ETC),
16 Throttle grip with grip travel sensor, 17 Knock sensor.



Cylinder charge

Component parts

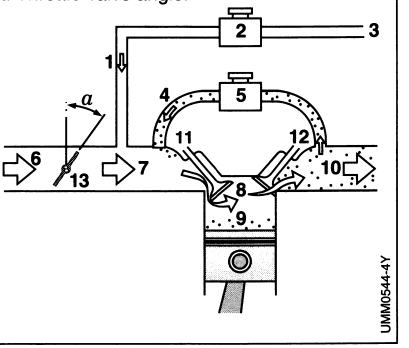
The gas mixture located in the cylinder after the intake valves have closed is termed the cylinder charge. It consists of the supplied fresh A/F mixture and the residual exhaust gas. The major components that influence the cylinder charge are shown in Figure 1. The throttle valve is the most important component on modern systems.

Fresh A/F mixture

The component parts of the fresh A/F mixture drawn in are fresh air and – on systems with external mixture formation – the fuel suspended in it (Figure 1). Most of the fresh air flows through the throttle valve; additional fresh A/F mixture can be drawn in through the evaporative-emissions control system. The air mass present in the cylinder after the intake valves have closed is the decisive factor in the work performed above the piston during combustion, and ultimately in the torque delivered by the engine. Therefore, measures for increasing maximum torque and maximum engine power almost always require an increase in the maximum possible cylinder charge. The theoretical maximum charge is predetermined by the piston displacement and, in the case of turbocharged engines, by the achievable charge-air pressure, as well.

Figure 1: Cylinder charge in a spark-ignition engine

1 Air and fuel vapor,
2 Canister-purge valve (tank vent valve),
3 Connection to evaporative-emissions control system,
4 Exhaust gas,
5 Exhaust-gas recirculation valve (EGR valve),
6 Air-mass flow (ambient pressure),
7 Air-mass flow (intake-manifold pressure),
8 Fresh A/F mixture charge (combustion-chamber pressure),
9 Residual exhaust-gas charge (combustion-chamber pressure),
10 Exhaust gas (exhaust-gas backpressure),
11 Intake valve,
12 Exhaust valve,
13 Throttle valve (butterfly valve).
α Throttle-valve angle.



Residual exhaust gas

The residual exhaust gas in the charge is formed by:

- the mass of exhaust gas that remains in the cylinder or is stored briefly during the charge cycle in the intake or exhaust port and then drawn back again, and is not ejected on completion of the charge cycle via the exhaust valve, and
- in the case of systems with exhaust-gas recirculation by the mass of the returned exhaust gas.

The amount of residual exhaust gas is determined by the charge cycle. It does not contribute directly to the combustion process, but does influence ignition and the combustion process as a whole. At wide-open throttle, the amount of residual exhaust gas generally needs to be as small as possible in order to maximize the fresh-air mass and also the power output of the engine.

At part load, however, the amount of residual exhaust gas is desirable in order to reduce fuel consumption. This is achieved by a more favorable cycle, resulting from a change in mixture composition, and also by a reduction in pump losses during the charge cycle, as a higher intake-manifold pressure is required for the same air charge. A specifically introduced amount of residual exhaust gas can likewise reduce the emission of nitrogen oxides (NO_x) and unburned hydrocarbons (HC).

Control of cylinder charge

On spark-ignition (SI) engines with external mixture formation (intake-manifold injection) and on systems with internal mixture formation (gasoline direct injection) and homogeneous cylinder charge, the torque delivered by the engine is determined by the air charge. In contrast to this, in the case of internal mixture formation with an air surplus, the engine torque can also be controlled directly by varying the injected fuel mass (stratified-charge operation).

Throttle valve

The critical adjustment mechanism for controlling air-mass flow supplied to the cylinders is the throttle valve. When the throttle valve is not fully open, the air flow drawn in by the engine is throttled, thus reducing maximum engine torque. This throttling effect is dependent on the position, and thus the opening cross-section of the throttle valve, as well as on engine speed (Figure 2). Maximum torque is developed with the throttle valve wide open.

By the end of the 1990s the mechanical throttle valve was state of the art. The driver directly controls the opening of the throttle valve via a Bowden cable by pressing the accelerator pedal.

On systems with electronic throttle control (ETC), the required air charge is calculated from the desired engine torque, which is obtained from the accelerator-pedal position, and the throttle valve is activated accordingly.

Electronic throttle control (ETC) components

The electronic throttle control system (Figure 3) consists of the accelerator-pedal module, the engine control unit, and the throttle device. The throttle device mainly consists of the throttle valve, the electric throttle-valve drive element, and the throttle-valve position sensor. The drive element is a DC motor, which acts on the throttle-valve shaft via a gear unit. The throttle-valve position sensor, designed as a redundant unit, detects the position of the throttle valve.

The driver command is detected by the redundant sensor system in the accelerator-pedal module, and the signal is sent to the engine control unit. This calculates the required cylinder charge, based on the current engine operating point, and regulates the opening angle of the throttle

Figure 2: Throttle map of a spark-ignition engine

- - - Intermediate position of the throttle valve.

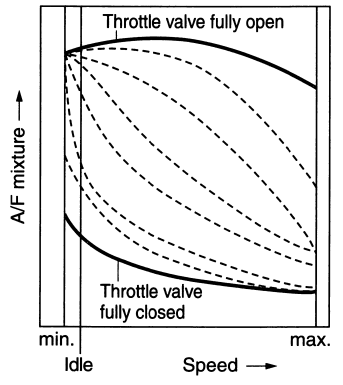
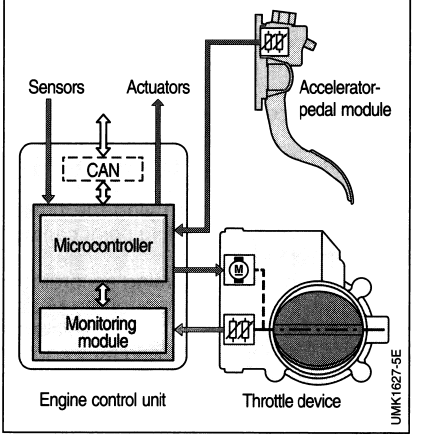


Figure 3: ETC system



valve by means of the throttle-valve drive element and the throttle-valve position sensor.

The redundancies incorporated in the accelerator-pedal module and the throttle device are part of the ETC monitoring concept for avoiding malfunctions.

Charge cycle

The charge cycle of fresh A/F mixture and residual exhaust gas is controlled by the opening and closing of the intake and exhaust valves. The opening and closing times of the valves (timing) and the valve lift curve are critical factors.

Valve timing

Valve overlap, i.e. the overlap of the opened times of the intake and exhaust valves, has a decisive influence on the mass of residual outlet gas remaining in the cylinder. The quantities of fresh A/F mixture and residual exhaust gas in the cylinder can be controlled by changing the valve-lift time curves.

In the overlap phase exhaust gas flows through the intake valve into the intake system due to the pressure differential between the exhaust system and the intake manifold. In the induction stroke it then returns to the combustion chamber (internal exhaust-gas recirculation). The longer the valve overlap lasts (early opening of the intake valve), the more mass of exhaust gas remains in the cylinder. This results in shortened combustion temperatures and thus a reduction in NO_x emissions.

However, since recirculated exhaust gas displaces fresh A/F mixture, opening the intake valve early also leads to a reduction in the maximum torque. In addition, excessively high exhaust-gas recirculation, particularly at idle, can lead to combustion misses which, in turn, cause an increase in HC emissions. With variable valve timing, varying the timing dependent on the operating point can optimize emissions.

By varying the valve timing accordingly (e.g. continuous adjustment of the phase position and valve lift), it is possible to control the air-mass flow, and thus engine performance, without having to use the throttle valve. The amount of residual exhaust gas can also be adjusted by means of valve timing.

On today's systems, the valves are actuated mechanically via the camshaft. Additional measures can be used to vary this activation to a certain degree (e.g. camshaft adjustment or camshaft shifting). However, these mechanical systems cannot dispense with the throttle valve completely.

Scavenging

In modern turbocharged engines valve-gear variabilities are additionally used to increase significantly the torque in the low-rpm range. This timing strategy is called scavenging. For this purpose the phase position of the intake valves and if necessary the exhaust valves are adjusted in the full-load range in such a way as to produce a valve overlap. During the valve overlap a partial flow of air passes directly again through the exhaust valve to the exhaust side, improving the scavenging of residual gas and with it enhancing anti-knock performance. At the same time the additional scavenging air improves process control at the turbocharger so that higher charge-air pressures can be achieved.

Exhaust-gas recirculation (EGR)

The mass of residual exhaust gas in the cylinder can also be increased by "external exhaust-gas recirculation" (EGR). In this case, an exhaust-gas recirculation valve connects the intake manifold and exhaust manifold (Figure 1). When the valve is open and with a pressure differential between the exhaust-gas system and the intake manifold, the engine draws in a combination of fresh A/F mixture and exhaust gas. The engine ECU calculates the level of exhaust-gas recirculation for a specific operating state, and activates the exhaust-gas recirculation valve accordingly.

Fuel reduction

Exhaust-gas recirculation increases intake-manifold pressure. The higher intake-manifold pressure leads to a reduction in the charge cycle work and this lowers fuel consumption.

NO_x reduction
Exhaust-gas recirculation is used on engines with gasoline direct injection which operate in lean-burn mode (stratified-charge operation) to reduce NO_x emissions. Exhaust-gas recirculation is the main way to minimize untreated NO_x emissions, and thus prolong the duration of lean-burn operation, which is limited by the NO_x storage catalytic converter.

Exhaust gas can be conducted back to the combustion chamber to reduce peak combustion temperatures. The reduction in temperature is based on the fact that the returned exhaust gas does not participate in the combustion and thus delivers no combustion energy. It is also an additional thermal mass, which means that the combustion energy distributes itself to a higher overall mass.

Higher combustion temperatures induce an over-proportional increase in NO_x formation, and as exhaust-gas recirculation (EGR) reduces combustion temperatures, it represents a particularly effective means of controlling NO_x emissions.

Supercharging
The torque obtainable is proportional to the charge of fresh A/F mixture. This enables an increase in the maximum torque in that the air in the cylinder is compressed by means of dynamic charging, mechanical charging or exhaust turbocharging (see Supercharging).

Recording of air charge

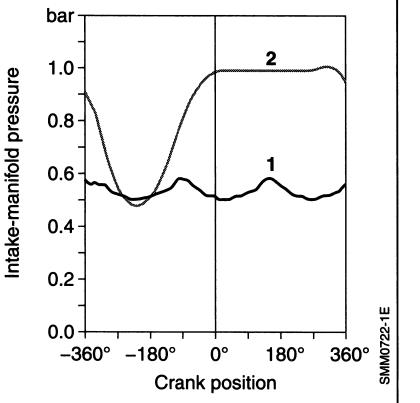
The cylinder charge crucially influences the power output, fuel consumption, and emissions. The composition of the air/fuel mixture plays a decisive role here (see Mixture formation, see Catalytic exhaust-gas treatment). To be able to adjust this correctly, it is necessary among others things to record the cylinder charge in the engine-management system with great precision. The throttle valve, as the central adjustment mechanism, determines the pressure in the intake manifold and uses this to influence the air-mass flow. Therefore the intake-manifold pressure is one of the central input variables which the engine management draws on to calculate the cylinder charge. The fresh-air mass however still depends on many additional factors, such as the ambient pressure, the throttle-valve leakage, the position of the gas-exchange valves, the operating status of the exhaust-gas turbocharger, the position of the swirl control valves, and the exhaust-gas recirculation. To take these influences into account, the engine-management system has in part further sensor technology, such as a hot-film air-mass meter to determine the cylinder charge.

Particularities for motorcycles
Motorcycles generally have less variability in the air system. Turbocharging, exhaust-gas recirculation and variable valve timing have barely featured up to now. Variable intake-manifold geometries have hitherto been the preserve of vehicles of higher performance classes. Particularly in small single-cylinder engines, therefore, the air charge is recorded purely by way of the throttle-valve position and intake-manifold pressure.

Compared with typical passenger-car applications, there are systematic differences in the intake manifold-pressure behavior. These are primarily attributable to different, geometric conditions. While in passenger-car applications a common throttle valve is for the most part used for a number of cylinders, motorcycles frequently rely on individual throttle valves (see Motorcycle intake manifolds). These are installed close to the cylinder head to guarantee good response. The absence of a classic manifold means that the volumes between the throttle valve and the intake valves are comparatively small. This results in small ratios of intake-manifold volume to displacement of the individual cylinder. This factor influences the intake manifold-pressure behavior (see Figure 4). While for a car a virtually constant vacuum is obtained in the intake manifold with a comparable load point and comparable engine speed, for a motorcycle the vacuum is achieved only briefly during the intake phase. As soon as the intake valves are closed, fresh air flows through the throttle valve and the intake-manifold pressure increases to ambient pressure. This behavior must be taken into account in the recording of charge by the engine management since the mean intake-manifold vacuum can only be used to a limited extent to determine the fresh-air mass.

Figure 4: Pressure characteristic in the intake manifold of car and motorcycle at part load

1 Three-cylinder car engine,
2 Single-cylinder motorcycle engine.

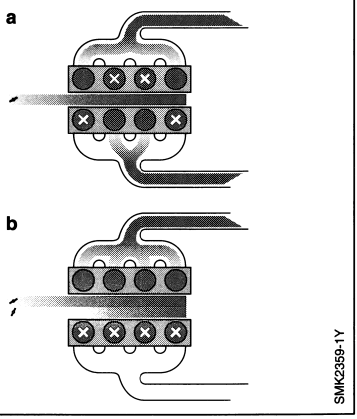


Cylinder shutoff

On engines with cylinder shutoff individual cylinders can be shut off at part load. The cylinders that continue to fire then generate the entire demanded torque. The air charge required for this is roughly doubled when every second cylinder is shut off referred to the corresponding full-engine operating point. The lower gas-exchange losses as a result of the de-throttling of the fired cylinders give rise to reduced fuel consumption. The gas-exchange losses can be further reduced by shutting off the gas-exchange valves of the non-fired cylinders. This shutoff necessitates accordingly designed valve gear (see Valve gear).

The most important motivation behind an engine design with cylinder shutoff is the reduction of fuel consumption. Since diesel engines are as a rule operated under de-throttled conditions to a large extent, the advantage is confined to the shutoff of the charge cycle. The reduced consumption that can be achieved by cylinder shutoff is dependent, aside from the operating-point profile, on the individual engine and drivetrain design. In gasoline engines with charge-cycle shutoff, it is realistic to achieve improvements in the higher single-digit percentage range in the conventional driving cycles and in the double-digit percentage range at individual operating points at low load.

Figure 5: Configuration of cylinder shutoff
a) Half-engine operation,
b) Cylinder-bank shutoff.



Configuration of cylinder shutoff

A cylinder shutoff can be implemented both in an engine with multiple cylinders and in an engine with only two cylinders. Different cylinder-shutoff configurations are possible here. The possible configurations are derived from the engine's design features. The most important criteria are the firing order and the respective layout of the intake and exhaust tracts. In most series applications every second cylinder in firing order and consequently half the cylinders is shut off. In half-engine operation (Figure 5a) shuttable-off and permanently fired cylinders are operated in the same bank. Here it is necessary to shut off the charge cycle, since λ control and the conversion capability of the catalytic converter would not be guaranteed with the oxygen that otherwise enters the exhaust tract.

When all the cylinders in an exhaust bank are shut off (cylinder-bank shutoff, Figure 5b), charge-cycle shutoff can also be dispensed with. This reduces the design complexity and the costs, but does not exploit the full CO_2 savings potential of cylinder shutoff.

Adaptation of vehicle components

With regard to NVH aspects (Noise Vibration Harshness), it may, aside from the valve gear, also be necessary to adapt the design of the dual-mass flywheel, the engine bearings, and the exhaust-gas system. An intake-manifold pressure sensor may also be needed to diagnose the valve actuators.

Engine management

Engine management is particularly important when it comes to operating an engine with cylinder shutoff. Partial-engine operation necessitates the adaptation of numerous software functions (Figure 6). Furthermore, additional functions for requirement and changeover coordination and diagnosis of the valve actuators for zero stroke are required.

Control of the changeover between full- and partial-engine operation is particularly complex because all the changeovers must take place imperceptibly to the driver and the response must not be delayed compared with full-engine operation. As well as charge adaption, exactly timed fading-out and -in of fuel injection for the shuttable-off cylinders and interventions via the moment of ignition are required for this purpose. Therefore, as well as engine coordination, torque structure, air, fuel, ignition and exhaust systems, and ETC monitoring are also adapted.

Cylinder-shutoff requirement

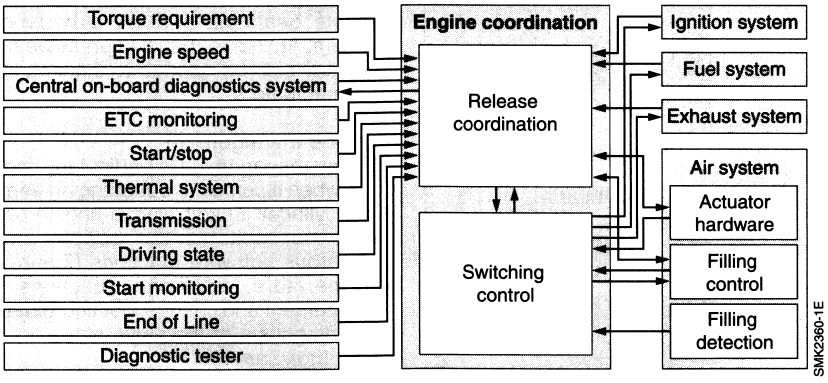
Operating-mode control by engine management constantly monitors numerous conditions for cylinder shutoff. Among the most important input variables is the setpoint torque to be developed by the engine, since the maximum settable torque is limited with cylinder shutoff. Another important input variable is the engine speed, since reliable switching of the gas-exchange valves is not possible above a limit speed. Further input variables are the engine temperature and the driving speed. Conditions are assigned to each input variable. On this basis either partial- or full-engine operation is requested.

Changeover between partial- and full-engine operation

The changeover to partial-engine operation begins with the increase of the air charge so that the still-fired cylinder can generate the entire demanded engine torque. In an engine with one intake manifold the increase is effected jointly for all the cylinders. In an engine with separate intake manifolds the air change can be changed for the still-fired cylinder bank regardless of the bank to be shut off. To deliver the required setpoint torque, the moment of ignition must be retarded accordingly. As soon as a sufficient torque reserve has been created by way of the air-charge buildup, the gas-exchange valves can be deactivated for the cylinders to be shut off and fuel injection faded out. The moment of ignition is then advanced again.

The changeover to full-engine operation begins with the activation of charge cycle and fuel injection and ignition-angle adjustment. The moment of ignition is advanced in parallel with the charge adaptation.

Figure 6: Overview of the coordination of engine management for cylinder shutoff



Fuel supply

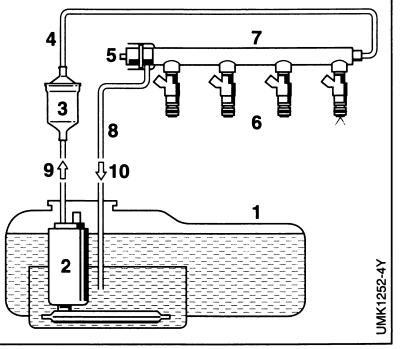
Fuel supply and delivery with intake-manifold injection

Different system configurations operating with intake-manifold injection typically at fuel pressures of approx. 300 to 400 kPa (3 to 4 bar) are used to supply fuel.

System with fuel return
The electric fuel pump delivers the fuel and generates the injection pressure (Figure 1). The fuel is drawn from the fuel tank and passes through the fuel filter into the high-pressure line, from where it flows to the engine-mounted fuel rail (fuel distributor). The rail supplies the fuel to the fuel injectors. The mechanical fuel-pressure regulator mounted on the rail keeps the differential pressure between the fuel injectors and the intake manifold constant, regardless of the absolute intake-manifold pressure, i.e. the engine load.

The fuel not needed by the engine flows through the rail via the return line connected to the fuel-pressure regulator back to the fuel tank. This excess fuel heated in the engine compartment causes the fuel temperature in the tank to rise. Fuel vapor is generated as a function of fuel temperature. Ensuring adherence to environmental-protection regulations, the vapors are routed through a tank-ventilation system for intermediate storage in a carbon canister until they can be returned through the intake manifold for combustion in the engine (see Evaporative-emissions control system).

Figure 1: Fuel supply and delivery with intake-manifold injection (system with fuel return)
1 Fuel tank, 2 Electric fuel pump, 3 Fuel filter, 4 Pressure line, 5 Fuel-pressure regulator, 6 Fuel injectors, 7 Rail (fuel flowing through), 8 Return line, 9, 10 Flow direction of fuel.



Returnless system
In a returnless fuel-supply system, the fuel-pressure regulator is located in the fuel tank or in its immediate vicinity, which means that the return line from the engine to the fuel tank is no longer required. Since the fuel-pressure regulator on account of its installation location has no reference to the intake-manifold pressure, the relative injection pressure is not dependent on the engine load. This is taken into account in the calculation of the injection duration in the engine control unit.

Only the amount of fuel which is to be injected is delivered to the rail. The excess flow volume delivered by the electric fuel pump returns directly to the fuel tank without taking the circuitous route through the engine compartment. In this way, fuel heating in the fuel tank and thus also evaporative emissions are significantly lower than in systems with fuel return. Because of these advantages, it is returnless systems which are predominantly used today.

Demand-regulated returnless system
In a demand-regulated system, the fuel-supply pump delivers only that amount of fuel that is currently used by the engine and that is required to set up the desired pressure. Pressure control is effected by means of a closed control loop in the engine control unit, whereby the current fuel pressure is recorded by a pressure sensor (Figure 2). This dispenses with the mechanical fuel-pressure regulator. To vary the delivery volume of the fuel-supply pump, its operating voltage is altered by means of a clock module that is triggered by the engine control unit.

The system is equipped with a pressure-relief valve to prevent the buildup of excessive pressure even during overrun fuel cutoff or after the engine has been switched off.

As a result of demand control, no excess fuel is compressed and thus the capacity of the electric fuel pump is minimized. This results in reduced fuel consumption compared with system with non-controlled electric fuel pumps. The in-tank fuel temperature can also be further reduced.

Further advantages of a demand-regulated system are derived from the variably adjustable fuel pressure. On the one hand, the pressure can be increased during hot starting to prevent the formation of vapor bubbles. On the other hand, it is possible above all in turbocharging applications to meter better the very large injected quantities required at high loads and the very small quantities at idle (injected-fuel-quantity spread) by effecting a pressure increase at full load and a pressure decrease at very low loads. This resolves the problem of metering small quantities in the fuel-injector configuration.

Furthermore, the measured fuel pressure provides for improved diagnostic options for the fuel system compared with previous systems. More precise fuel metering is achieved by taking into account the current fuel pressure when calculating the injection time.

Fuel supply and delivery with gasoline direct injection

Compared with injecting fuel into the intake manifold, there is only a limited time window available for injecting fuel directly into the combustion chamber. Mixture formation is therefore more important and fuel must be injected in the case of direct injection at a pressure approximately 50 times greater than is the case with intake-manifold injection.

The fuel system is divided into a low-pressure circuit and a high-pressure circuit.

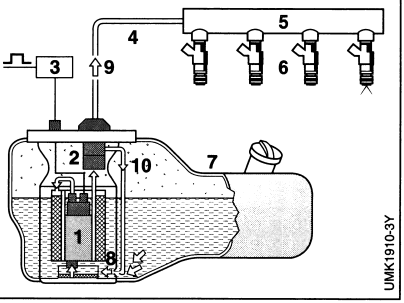
Low-pressure system

With gasoline direct injection the low-pressure system serves to supply the high-pressure system, involving the use of the same fuel systems and components as intake-manifold injection. Because of the high temperatures at the high-pressure pump in hot-start and hot-running conditions, a higher predelivery pressure (admission pressure) is required to prevent the formation of vapor bubbles. It is therefore an advantage to use systems with variable low pressure. Demand-regulated low-pressure systems are therefore particularly suitable here, because they allow the optimum admission pressure to be set for every engine operating state; this admission pressure usually varies in a range of 300 to 600 kPa (3 to 6 bar) relative to the ambient pressure.

Demand-controlled systems are increasingly being used as a further variant. In contrast to demand-regulated systems,

Figure 2: Fuel supply and delivery with intake-manifold injection (demand-regulated system)

1 Electric fuel pump with fuel filter (fuel filter may also be mounted outside of the fuel tank),
2 Pressure-relief valve and pressure sensor (pressure sensor alternatively in the rail),
3 Clock module for controlling el. fuel pump,
4 Pressure line, 5 Rail (returnless),
6 Fuel injectors, 7 Fuel tank,
8 Suction-jet pump,
9, 10 Flow direction of fuel.



in which the pressure is adjusted via a closed control loop by the pressure sensor, the pressure sensor is omitted here. The electric fuel pump is activated 100 % by pilot control as a function of the engine operating point.

High-pressure system

High-pressure fuel circuits use predominantly demand-regulated high-pressure pumps as well as continuous-delivery high-pressure pumps. The system comprises the fuel rail (high-pressure rail) with high-pressure injectors and a high-pressure sensor (Figures 3 and Figure 4). The continuous-delivery system also requires a separate pressure-control valve.

Depending on the engine's operating point, the pressure varies typically in a continuous-supply system in a range of 5 to max. 11 MPa (50 to 110 bar), and in a demand-regulated system up to max. 35 MPa (350 bar). The high-pressure sensor provides the information on the currently prevailing pressure.

The signal from the high-pressure sensor is also used for injection calculation and for fuel-system diagnostics.

Continuous-delivery system

The high-pressure pump driven by the engine camshaft, normally a three-barrel radial-piston pump (see High-pressure pumps for gasoline direct injection), forces fuel into the rail against the system pressure (Figure 3). The pump delivery quantity cannot be adjusted. The excess fuel not required for fuel injection and to maintain the pressure is depressurized by the pressure-control valve and returned to the low-pressure circuit. For this purpose, the pressure-control valve is actuated by

Figure 3: Fuel supply and delivery with gasoline direct injection (continuous-delivery system)

1 Suction-jet pump,
2 Electric fuel pump with fuel filter,
3 Pressure regulator,
4 Low-pressure line,
5 Continuous-delivery high-pressure pump,
6 High-pressure line,
7 High-pressure sensor,
8 Rail,
9 Pressure-control valve,
10 High-pressure injectors
11, 12 Flow direction of fuel.

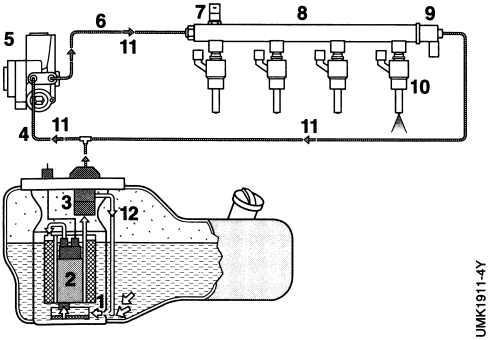
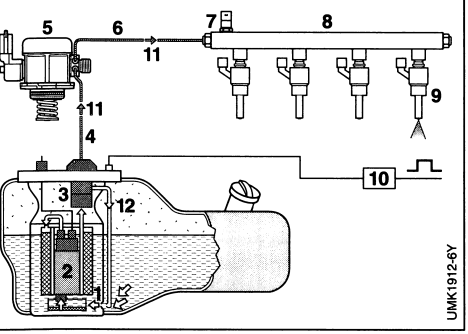


Figure 4: Fuel supply and delivery with gasoline direct injection (demand-regulated system)

1 Suction jet pump,
2 Electric fuel pump with fuel filter,
3 Pressure-relief valve and pressure sensor (pressure sensor alternatively in the low-pressure line),
4 Low-pressure line,
5 Demand-regulated high-pressure pump with integrated fuel-supply control valve and pressure-relief valve,
6 High-pressure line,
7 High-pressure sensor,
8 Rail,
9 High-pressure injectors,
10 Clock module for controlling the electric fuel pump,
11, 12 Flow direction of fuel.



the engine control unit in such a way as to obtain the injection pressure required at a given operating point. The pressure-control valve also serves as a mechanical pressure-relief valve.

In continuous-delivery systems, most of the operating points cause significantly more fuel to be compressed to high system pressure than is needed by the engine. This results in unnecessary energy consumption and thus higher fuel consumption than demand-regulated systems. The excess fuel released via the pressure-control valve also contributes to increasing the temperature in the fuel system. For these reasons, only demand-regulated high-pressure systems are used in modern direct-injection engines.

Demand-regulated system
In a demand-regulated system (Figure 4), the high-pressure pump – usually a single-barrel radial-piston pump (see High-pressure pumps for gasoline direct injection) – delivers to the fuel rail only the precise amount of fuel which is actually needed for injection and to set the desired pressure. The pump is usually driven by the engine camshaft, single-barrel pumps being driven via special cams which drive the pump plunger. The delivery quantity is varied by a fuel-supply control valve which is integrated in the high-pressure pump. The engine control unit actuates this valve with each pump lift with such precision as to provide the required delivery quantity in order to set the system pressure needed in the rail for the given operating point.

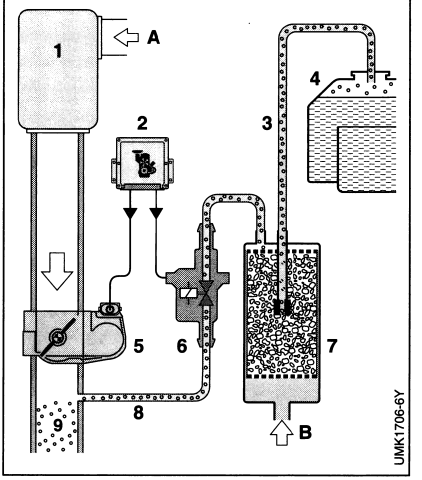
For safety reasons the high-pressure circuit incorporates a mechanical pressure-relief valve, which is usually integrated directly in the high-pressure pump. Should the pressure exceed the permissible level, fuel is returned via the pressure-relief valve to the low-pressure circuit.

Evaporative-emissions control system

Function

An evaporative-emissions control system (tank ventilation) is required for vehicles equipped with a spark-ignition (SI) engine. Its purpose is to intercept and collect fuel degassing from the fuel tank and to comply with the legal stipulations governing emission limits for evaporative loss. Degassing increases when the fuel in the fuel tank is heated. This may be caused by the higher ambient temperatures, the power loss of the fuel pump integrated in the fuel tank or – depending on the fuel supply system – the return of fuel heated in the engine and no longer required for the combustion process. Degassing also occurs when the ambient pressure drops,

Figure 5: Evaporative-emissions control system
1 Air filter,
2 Engine control unit,
3 Fuel-tank vent line,
4 Fuel tank,
5 Throttle device,
6 Canister-purge valve,
7 Carbon canister,
8 Line to intake manifold,
9 Intake manifold.
A Intake air,
B Fresh air.



for example due to weather influences or when driving uphill.

Design and operating principle

The evaporative-emissions control system consists of a carbon canister, into which the vent line from the fuel tank projects, as well as a canister-purge valve that is connected both to the carbon canister as well as to the intake manifold (Figure 5). The activated carbon absorbs the fuel vapor. Due to the vacuum prevailing in the intake manifold, fresh air is drawn through the activated carbon when the canister-purge valve frees up the line between the carbon canister and the intake manifold when driving. The fresh air picks up the absorbed fuel and feeds it to the combustion process. This is know as purging the carbon canister.

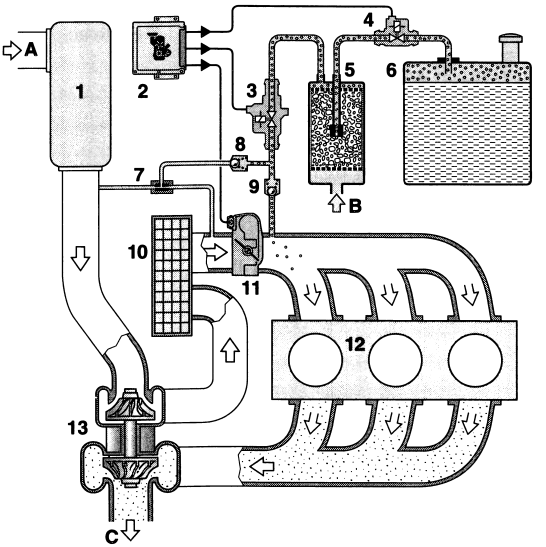
The engine control unit controls the purge-gas volume depending on the engine operating point. In order to ensure that the carbon canister is always able to absorb fuel vapor, the activated carbon must be regenerated at regular intervals. Since current spark-ignition engines are operated to increase efficiency in such a way that a high intake-manifold vacuum is avoided (e.g. with downsizing, turbo operation, lean-burn concepts), adequate regeneration is often not feasible. In this case an additional regenerating line is used which utilizes the pressure drop after the air filter for additional vacuum generation (Figure 6). In turbocharged engines a venturi tube can also be used for this purpose. A non-return valve is integrated in each ventilation path so that there are no circular flows.

Pressure tank

With hybrid vehicles, which can be run on purely electric power for longer distances, there is in some cases no further opportunity for regeneration. In this case, the usual fuel tank made of plastic is replaced by a pressure-resistant steel tank which is connected via a further shutoff valve mounted on the tank to the carbon canister (Figure 6). In the event of increased degassing the pressure in such a tank rises, which counteracts further degassing of the fuel. If the pressure in the tank

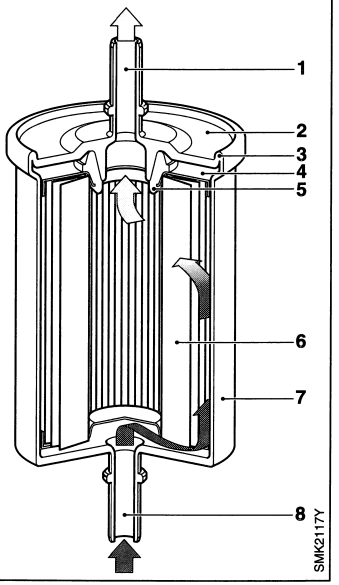
Figure 6: Evaporative-emissions control system for hybrid vehicles

1 Air filter,
2 Engine control unit,
3 Canister-purge valve,
4 Shutoff valve,
5 Carbon canister,
6 Fuel tank (pressure-resistant steel tank),
7 Venturi tube,
8 Non-return valve,
9 Non-return valve,
10 Intercooler,
11 Throttle device,
12 Engine,
13 Exhaust-gas turbocharger.
A Intake air,
B Fresh air,
C Exhaust gas.



reaches critical values, the tank pressure is regulated by the engine ECU, whereby the additional shutoff valve opens and the fuel vapors are channeled into the carbon canister. The tank pressure must also be reduced to refuel the vehicle. For regeneration, operation of the internal-combustion engine is then explicitly requested by the engine ECU.

Figure 7: In-line gasoline filter
1 Fuel outlet, 2 Filter cover,
3 Laser-welded edge, 4 Support ring,
5 Seal,
6 Spiral vee-shaped filter medium,
7 Pressure-stable filter housing,
8 Fuel inlet.



Gasoline filter

Function

The function of the fuel filter is to filter the fuel that flows into the fuel system. Contaminants in the fuel must be filtered out in order to protect the system and in particular the fuel injectors.

Design

Fuel filters for spark-ignition engines (gasoline filters) are located on the pressure side after the fuel-supply pump. In-tank filters are the preferred choice in newer vehicles, i.e. the filter is integrated in the fuel-supply module in the fuel tank. In this case, it must always be designed as a lifetime filter, which does not need to be changed over the full service life of the vehicle. Furthermore, in-line filters, which are installed in the fuel line, continue to be used. These can be designed as replacement parts or lifetime parts.

The filter housing is manufactured from steel, aluminum or plastic. It is connected to the fuel feed line by a thread, tube or quick-action connection. The housing contains the filter element, which filters the dirt particles out of the fuel (Figure 7). The filter element is integrated in the fuel circuit in such a way that fuel passes through the entire surface of the filter medium as much as possible at the same flow velocity.

Filter medium

Special resin-impregnated cellulose-fiber papers which are also bonded for heavier-duty applications to a synthetic-fiber (meltblown) layer are used as the filter medium. This bond must ensure high mechanical, thermal and chemical stability. The paper porosity and the pore distribution of the filter paper determine the filtration efficiency and throughflow resistance of the filter.

Fuel filters for gasoline engines are either spiral vee-form or radial vee-form in design. In a spiral vee-form filter, an embossed filter paper is wrapped round a support tube. The fuel flows through the filter in the longitudinal direction.

In a radial vee-form filter, the filter paper is folded and inserted into the housing in the shape of a star (Figure 7). Plastic,

resin or metal end rings and, if necessary, an inner protective jacket provide stability. The fuel flows through the filter from the outside inwards, during which the dirt particles are separated from the filter medium.

Requirements

The required filter fineness is dependent on the fuel-injection system. For systems with intake-manifold injection, the filter element has a mean pore width of approx. 10 µm. Finer filtering is required for gasoline direct injection. The mean pore width here is in the range of 5 µm. Particles larger than 5 µm must be separated off by up to 85 %. In addition, a filter for gasoline direct injection, when new, must satisfy the following residual-dirt requirement: Metal, mineral and plastic particles and glass fibers with diameters of more than 400 µm must not be flushed by the fuel out of the filter.

Filter efficiency depends on the throughflow direction. When replacing in-line filters, it is imperative that the flow direction specified by the arrow be observed.

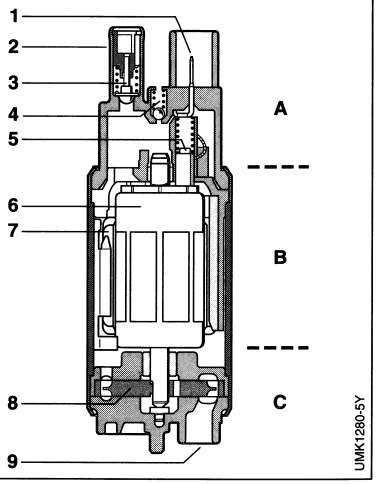
The interval for changing conventional in-line filters is, depending on filter volume and fuel contamination, normally between 30,000 km and 120,000 km. In-tank filters can generally be used for at least 250,000 km before they have to be replaced. This corresponds to today's design lifetime for a spark-ignition engine. There are in-tank and in-line filters available for use with gasoline direct-injection systems which feature service lives in excess of 250,000 km. If the in-tank filter has to be exchanged, this is only possible by completely changing the fuel-supply module.

Electric fuel pump

The electric fuel pump must deliver sufficient quantities of fuel to the engine and maintain the pressure required for efficient fuel injection in all operating states. Essential requirements include:
- Delivery quantity between 40 and 300 *l*/h at nominal voltage
- Pressure in the fuel system between 300 and 650 kPa (3.0 to 6.5 bar)
- Buildup of system pressure from 50 to 60 % of nominal voltage; the decisive factor here is operation during cold starting.

Apart from this, the electric fuel pump is increasingly being used as the pre-supply pump for the modern direct-injection systems used on diesel and gasoline engines. In the case of gasoline direct injection, sometimes pressures of up to 650 kPa must be provided during hot-delivery operation.

Figure 7: Electric fuel-pump design using a flow-type pump as an example
A Fitting cover,
B Electric motor,
C Pump barrel.
1 Electrical connection,
2 Hydraulic connection (fuel outlet),
3 Non-return valve,
4 Pressure-relief valve
5 Commutator with carbon brushes,
6 Motor armature with windings which are energized by the commutator,
7 Magnet,
8 Impeller ring of flow-type pump,
9 Hydraulic connection (fuel inlet).



Electric motor
The plunger-and-barrel assembly (pump element) of the electric fuel pump (Figure 7) is driven by an electric motor. An armature with copper or carbon commutator is used as standard in this engine. Electronic commutating systems without commutators and carbon brushes are increasingly being used in new vehicles that appear on the market. The design of the electric motor depends on the desired delivery quantity for a given system pressure. It is continuously surrounded by flowing fuel and thereby cooled.

End cover
The end cover contains the electrical connection and the pressure-side hydraulic connection. The non-return valve prevents the fuel lines from draining after the electric fuel pump is switched off; this ensures that the system pressure is maintained for a certain period of time after switch-off.

A pressure-relief valve is integrated as required. Most end covers also include the carbon brushes for the commutation system and interference-suppression elements (inductance coils, with capacitors in some applications).

Plunger-and-barrel assembly
The plunger-and-barrel assembly is designed as a positive-displacement or flow-type pump.

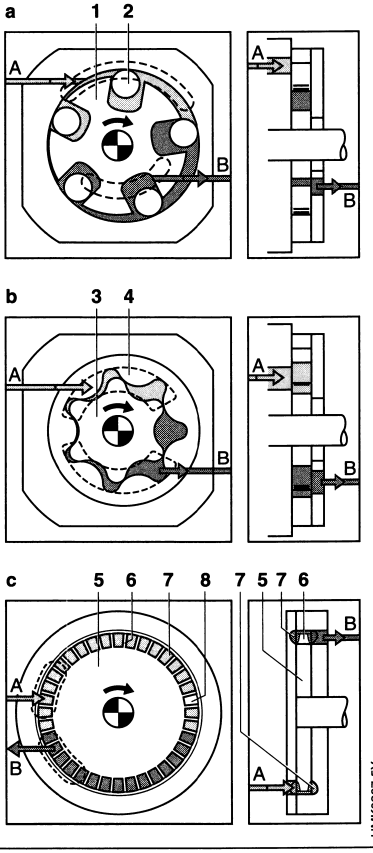
Positive-displacement pump
In a positive-displacement pump, volumes of liquid are basically drawn in and transported in a closed chamber (apart from leaks) by rotation of the pump element to the high-pressure side. A roller-cell pump (Figure 8a), an internal-gear pump (Figure 8b) or a screw-spindle pump may be used for the electric fuel pump.

Roller-cell pump
The slotted rotor arranged eccentrically in the pump housing has metal rollers around its circumferences which are

loosely guided in the slot-shaped openings. The centrifugal force generated by the rotation of the slotted rotor and the fuel pressure press the rollers against the pump housing and the driving flanks of the slots. The rollers operate here as circulating seals. The fuel is delivered in the chamber created between in each case two rollers of the slotted rotor and the pump housing. The pumping effect is created by the fact that, after the inlet opening is sealed, the chamber volume is continuously reduced until the fuel leaves the pump through the outlet opening.

Figure 8: Electric fuel-pump principles
a) Roller-cell pump,
b) Internal-gear pump,
c) Flow-type pump.

A Fuel inlet (suction opening),
B Outlet (pressure side).
1 Slotted rotor (eccentric),
2 Roller,
3 Internal drive wheel,
4 Rotor (eccentric),
5 Impeller,
6 Impeller blades,
7 Passage (peripheral),
8 Gas-discharge orifice.



Internal-gear pump
The internal-gear pump consists of an internal drive wheel which engages an eccentrically arranged outer rotor. This outer rotor has one tooth more than the drive wheel. The tooth flanks that are sealed against each other create during the rotation process in their spaces variable chambers which provide the pumping effect.

Positive-displacement pumps are advantageous when used with high-viscosity media, e.g. cold diesel fuel. They have a good low-voltage characteristic, i.e. they have a relatively "flat" delivery-rate characteristic over the operating voltage. Efficiency can be as high as 40 %. The unavoidable pressure pulses may cause noise; the extent of this problem varies according to the pump's design configuration and mounting location.

Whereas in electronic gasoline-injection systems the positive-displacement pump has to a large extent been superseded by the flow-type pump for the classical electric-fuel-pump requirements, it has gained a new field of application as the pre-supply pump on diesel common-rail systems with their significantly greater pressure requirements and viscosity range.

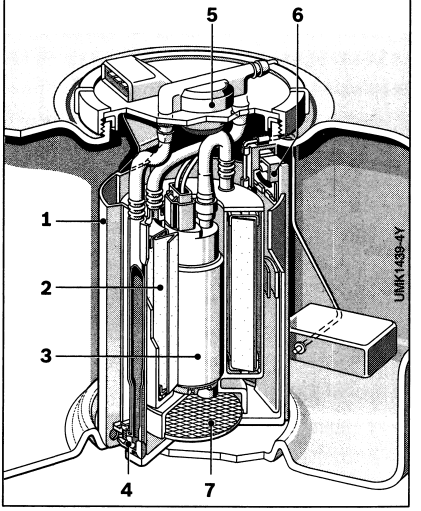
Flow-type pump
The flow-type pump has become the accepted solution for gasoline applications. An impeller ring equipped with numerous peripheral vanes rotates within a chamber consisting of two fixed housing sections (Figure 8c). Each of these sections features a passage along the path of the impeller blades, with openings at one end of the passage starting at the suction openings. From there, they extend to the point where the fuel exits

the plunger-and-barrel assembly at system pressure. A small gas-discharge orifice, located at a specified angular distance from the suction opening, improves performance when pumping hot fuel; this orifice facilitates the discharge of any gas bubbles that may have formed (with minimal leakage).

Pressure builds up along the passage as a result of the exchange of pulses between the impeller blades and the liquid particles. This leads to spiral-shaped rotation of the liquid volume trapped in the impeller and in the passages.

Flow-type pumps feature a low noise level since pressure buildup takes place continuously and is practically pulsation-free. Pump design is also much less complex than the positive-displacement unit. System pressures of up to 650 kPa can be achieved. The efficiency of these pumps can, depending on the operating point, exceed 30 %.

Figure 9: Fuel-supply module
1 Swirl pot,
2 Fuel filter,
3 Electric fuel pump,
4 Suction jet pump,
5 Fuel-pressure regulator,
6 Fuel-level sensor,
7 Prefilter.



Fuel-supply module

Whereas in the early stages of electronic gasoline injection the electric fuel pump was mounted exclusively outside the fuel tank (in-line), it is common practice today to install the pump inside the tank itself. The electric fuel pump forms an integral part of a fuel-supply module (in-tank unit, Figure 9) which may comprise further elements:

- a swirl pot,
- a fuel-level sensor (see Sensors),
- a fuel-pressure regulator for returnless fuel supply systems,
- a prefilter for protecting the pump,
- a pressure-side fine-mesh fuel filter, which will not need replacing for the entire life of the vehicle,
- and electrical and hydraulic connections.
- Tank pressure sensors (for diagnosing tank leaks), fuel-pressure sensors (for demand-regulated systems) and valves for tank ventilation may also be incorporated.

The swirl pot is situated centrally in the fuel tank. The suction side of the electric fuel pump is inserted in this beaker-shaped container. When the tank level is high, fuel flows over the beaker edge, when the level is low, fuel passes through a small opening in the beaker base or through the return line of the electric fuel pump into the swirl pot. When the vehicle is cornering, accelerating and braking, the fuel in the fuel tank flows to one side, which, when the tank level is low, could cause the fuel supply to be interrupted. The fuel can only escape slowly from the swirl pot through the small opening such that the fuel supply is safeguarded for a certain period of time.

The swirl pot can be filled passively, for example by a flap system or a change-over valve, or actively by a suction jet pump. This pump is designed as a venturi nozzle and does not contain any moving parts. The returning fuel flows as a driving medium into this pump and is accelerated after being discharged through the nozzle. In accordance with Bernoulli's law this produces a pressure drop. In this way, fuel from the fuel tank is drawn in through a second inlet and delivered under pressure with the returning fuel to the pump outlet.

High-pressure pumps for gasoline direct injection

The function of the high-pressure pump is to compress the fuel supplied in sufficient quantity by the electric fuel pump at an rpm- and temperature-dependent admission pressure (program-map value) to the level required for high-pressure injection.

Demand-regulated high-pressure pump

Design and operating principle

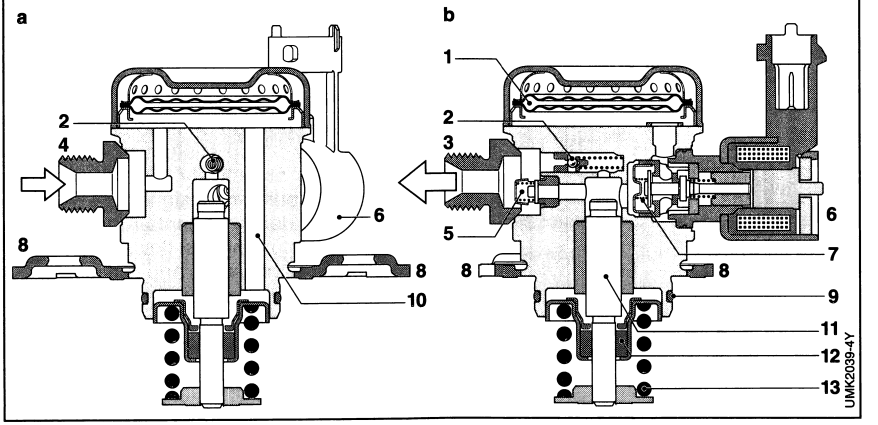
The Bosch demand-regulated high-pressure pump is used in systems for gasoline direct injection for injection pressures up to 20 MPa (2nd-generation gasoline direct injection). Subsequent further development of the second generation increases the pressure range to 25 MPa and improves the acoustics of the fuel-supply control valve. The third generation increases the pressure range to 35 MPa, facilitates a greater delivery quantity, and once again improves the acoustics.

The high-pressure pump is a cam-driven single-barrel pump running in oil (Figure 10) with an integrated fuel-supply control valve (also called the metering unit), an integrated pressure-relief valve on the high-pressure side and an integrated fuel-pressure attenuator on the low-pressure side. To comply with future stringent fuel standards and in response to emission-control legislation, the pump is made of stainless steel and is welded at all the emission-relevant connection points.

The high-pressure pump is mounted as a plug-in pump on the cylinder head. The interface between the engine camshaft and the delivery barrel is a bucket tappet for a double cam (Figure 11a), and a roller tappet for triple and quadruple cams (Figure 11b). This ensures that the cam lift curve is transferred to the delivery plunger of the high-pressure pump with the requirements with regard to lubrication, Hertzian stress and mass inertia. During the cam lift the tappet follows the contour of the cam. This results in the

Figure 10: Demand-regulated single-cylinder high-pressure pump for 2nd-generation gasoline direct-injection systems

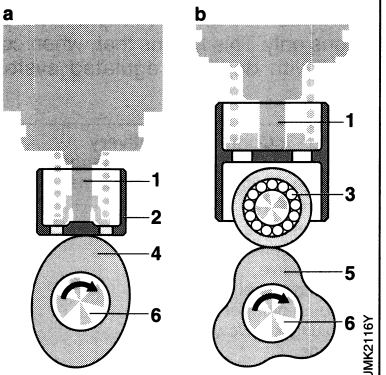
a) View with low-pressure port,
b) View with high-pressure port (on the same plane offset at an angle to low-pressure port).
1 Variable fuel-pressure attenuator, 2 Pressure-relief valve, 3 High-pressure port,
4 Low-pressure port, 5 Outlet valve,
6 Fuel-supply control valve, 7 Inlet valve,
8 Mounting flange, 9 O-ring,
10 Connection to plunger step space (pressure-attenuation function),
11 Delivery plunger, 12 Plunger seal, 13 Plunger spring.



vertical motion, or stroke, of the delivery plunger. In the delivery stroke the push rod absorbs the applied pressure, mass, spring and contact forces. It is rotationally secured in the process.

With a quadruple cam, time synchronization of delivery and injection is possible in a four-cylinder engine, i.e. each injection is also accompanied by a delivery. In this way, it is possible – due to the pressure dip in the rail – on the one hand to reduce excitation of the high-pressure circuit and on the other hand to reduce the rail volume.

Figure 11: High-pressure pump drives
a) Drive via bucket tappet,
b) Drive via roller tappet.
1 Pump plunger,
2 Bucket tappet,
3 Roller tappet,
4 Double cam,
5 Triple cam,
6 Drive shaft (engine camshaft).

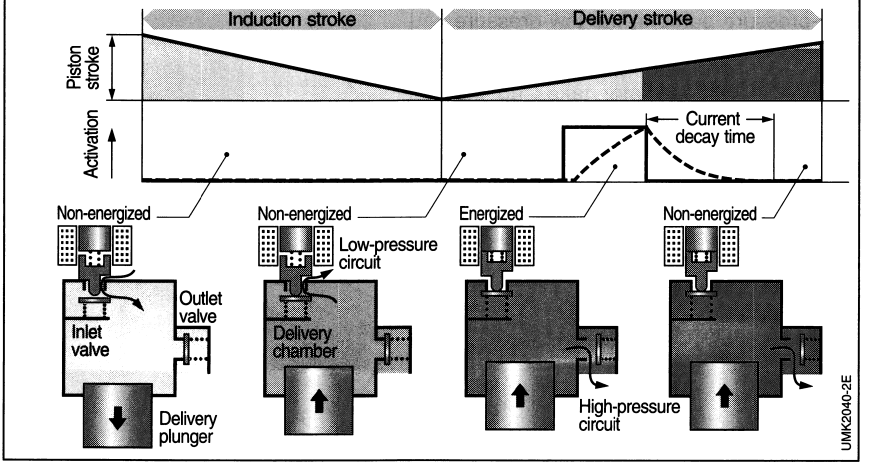


In order to ensure that the system pressure can still be sufficiently quickly achieved or varied with the engine at maximum fuel demand, the maximum delivery quantity of the high-pressure pump is configured for maximum demand plus necessary factors which affect delivery behavior (e.g. high-pressure starting, hot gasoline, pump aging, dynamic response).

The volumetric efficiency is derived from the ratio of actually delivered fuel quantity to theoretically possible quantity. This is dependent on the delivery-plunger diameter and stroke. The volumetric efficiency is not constant across all engine speeds. It is dependent on the following factors:

- In the lower speed range: plunger and other leakages.
- In the upper engine speed range: dynamics (inertia and opening pressure) of the inlet and outlet valves.
- In the total speed range: dead volume of the delivery chamber and temperature dependence of fuel compressibility.

Figure 12: Activation concept of fuel-supply control valve



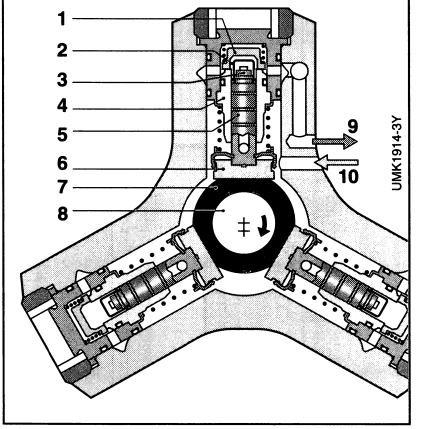
Fuel-supply control valve
Demand control of the high-pressure pump is effected with the fuel-supply control valve (Figure 12). The fuel delivered by the electric fuel pump is drawn into the delivery chamber via the inlet valve of the open fuel-supply control valve. In the subsequent delivery stroke, the fuel-supply control valve remains open after bottom dead center so that fuel that is not needed at the relevant load point is returned at admission pressure to the low-pressure circuit. After the fuel-supply control valve is actuated, the inlet valve closes, the fuel is compressed by the pump plunger and delivered to the high-pressure circuit. The inlet valve then remains closed during the delivery stroke on account of the pressure in the delivery chamber even when the fuel-supply control valve is non-energized. The engine-management system calculates the time from which the fuel-supply control valve is actuated as a function of delivery quantity and rail pressure. The start of delivery is varied here for demand control.

Fuel-pressure attenuator
The variable fuel-pressure attenuator (Figure 10) integrated in the high-pressure pump, together with the delivery plunger designed as a stepped plunger (backpumping in the induction stroke due to the smaller diameter in the lower area causes a repumping effect), attenuate the pressure pulsations stimulated by the high-pressure pump in the low-pressure circuit and at high rotational speeds ensure a good delivery-chamber charge. The pressure attenuator takes up the fuel quantity diverted at the relevant operating point via the deformation of its gas-filled metal diaphragms and releases it again in the induction stroke to fill the delivery chamber. Operation with variable admission pressure – i.e. the use of demand-regulated low-pressure systems – is possible here.

Continuous-delivery high-pressure pump
Design and operating principle
The Bosch continuous-delivery high-pressure pump is used in systems for gasoline direct injection for injection pressures up to 12 MPa (1st-generation gasoline direct injection). It is a radial-piston pump with three delivery barrels situated at circumferential offsets of 120° (triple-barrel radial-piston pump, Figure 13).

The delivery quantity of the continuous-delivery high-pressure pump is proportional to the rotational speed. The three barrels deliver fuel at circumferential offsets of 120° in order to ensure overlapping and therefore continuous delivery. This gives rise to minimal pressure pulsations only. This means that, when compared with demand-regulated systems

Figure 13: Continuous-delivery high-pressure pump for 1st-generation gasoline direct-injection systems (cross-section, schematic)
1 Outlet valve,
2 Displacement space,
3 Inlet valve,
4 Pump barrel,
5 Pump plunger (plunger hollow, fuel inlet),
6 Shoe,
7 Lift ring,
8 Eccentric element,
9 High-pressure port to rail,
10 Fuel supply (low-pressure).



with single-plunger pumps, less demands have to be placed on the pump connections and piping. Furthermore, there is no need for a low-pressure attenuator. On the downside, continuous delivery at high pressure means that it has a higher power loss than a demand-regulated pump.

When the pump operates at constant rail pressure or at part load, the pressure of the excess fuel is reduced to admission pressure level by the pressure-control valve mounted on the rail and returned to the intake side of the high-pressure pump. The pressure level in the high-pressure circuit is regulated and adjusted by means of the engine control unit, which specifically actuates the pressure-control valve.

Pressure-control valve

The pressure-control valve (Figure 14) is a proportional control valve which is closed at zero current and actuated by means of a pulse-width-modulated signal. During operation, the energizing of the solenoid coil sets a magnetic force which relieves the load on the spring, lifts the valve ball off the valve seat, and thereby alters the flow cross-section. The pressure-control valve sets the desired rail pressure as a function of the pulse duty factor. The excess fuel delivered by the high-pressure pump is diverted into the low-pressure circuit.

A pressure-limiting function is integrated by way of the valve spring to protect the components against unacceptably high rail pressures, for example, in the event of actuation failure. If one or more of the pump barrels should fail, emergency operation is possible with the intact barrels or by means of the electric fuel pump with admission pressure.

Fuel rail

Intake-manifold injection

The function of the fuel rail is to store the fuel required for injection, to attenuate pulsations and to ensure uniform distribution to all the fuel injectors. The fuel injectors are mounted directly on the rail. Next to the fuel injectors it is possible in systems with fuel return to integrate a fuel-pressure regulator and if necessary a fuel-pressure attenuator in the rail.

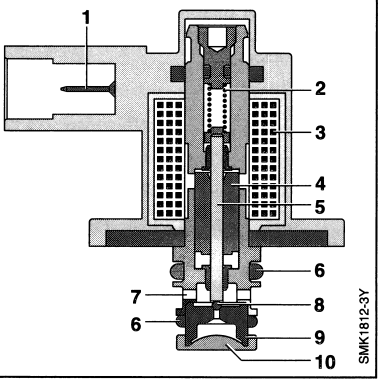
Local pressure fluctuations caused by resonance when the fuel injectors open and close is prevented by careful selection of the fuel-rail dimensions. As a result, irregularities in injected fuel quantity which can arise as a function of load and engine speed are avoided.

Gasoline direct injection

The fuel rail (Rail, Figure 15) has the task of storing and distributing the volume of fuel required for each operating point. The fuel is stored by way of its volume and compressibility. In this way, the volume or the inside diameter is engine-power-dependent and must be adapted to the relevant engine demand (injected fuel quantity) and pressure range. Full design of the rail is provided by the dimensioning

Figure 14: Pressure-control valve

1 Electrical connection, 2 Valve spring,
3 Solenoid coil, 4 Solenoid armature,
5 Valve needle, 6 Sealing ring (O-rings),
7 Drain hole, 8 Valve ball,
9 Valve seat, 10 Inlet with inlet strainer.



of the outside diameter or the wall thickness of the rail to guarantee strength. The rail is designed to be operationally stable, which is checked by way of special pressure load profiles which have been derived from real operating loads. The volume also ensures hydraulic attenuation in the high-pressure range, i.e. pressure fluctuations in the rail are compensated.

The rail consists of different add-on components which act either as interfaces or as functional groups. Specifically, these are:

- a stainless-steel pipe for accommodating the volume and mounting further attachments (copper-brazed or screwed),
- retainer for securing the rail to the engine block,
- cup for accommodating the fuel injectors,
- high-pressure line connection with integrated restrictor,
- pressure-sensor connection,
- end cap,
- hold-down element for securing the injector position during operation with combustion-chamber pressure and vibration acceleration,
- wiring harness with clipped-on plastic duct (around the pipe).

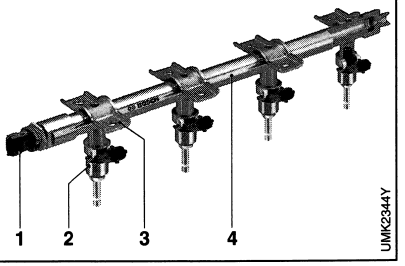
In part these solid, constructed or deep-drawn attachments are also designed as function modules (e.g. retainer and cup). The fuel injectors are sealed by O-rings, the pressure sensor and the high-pressure connection by a screwed ball-taper geometry.

A particular feature is the "suspended design" of the injector mounting, which ensures a low load thanks to its pressure-compensated construction. The drawback of this variant is the increased amount of space required.

The biggest challenge in the development of rails is the large number of necessary solutions in that it must always be integrated in the engine in an already specified installation space.

The rail for 1st-generation fuel-injection systems is designed for a pressure range up to 12 MPa (plus 0.5 MPa pressure-relief valve opening pressure). For the 2nd generation the pressure range extends up to 25 MPa (plus 5.0 MPa pressure-relief valve opening pressure), for the 3rd generation according to the high-pressure pump up to 35 MPa. The acoustics of the rail in the overall system are also improved as part of the further development.

Figure 15: For gasoline direct injection
1 High-pressure sensor,
2 Fuel injectors,
3 Retainer,
4 Rail.



Fuel-pressure regulator

Function

With intake-manifold injection the amount of fuel injected by the fuel injector depends on the injection period and the pressure differential between the fuel pressure in the fuel rail and the back pressure in the intake manifold. On fuel systems with fuel return, the influence of pressure is compensated for by a fuel-pressure regulator which maintains the differential between fuel pressure and intake-manifold pressure at a constant level. This pressure regulator permits just enough fuel to return to the fuel tank so that the pressure drop across the fuel injectors remains constant. In order to ensure that the fuel rail is efficiently flushed, the fuel-pressure regulator is normally located at the end of the rail.

On returnless fuel systems, the fuel-pressure regulator is part of the fuel-supply module in the fuel tank. The fuel-rail pressure is maintained at a constant level with reference to the ambient pressure. This means that the pressure differential between fuel-rail pressure and intake-manifold pressure is not constant and must be taken into account when the injection duration is calculated.

Mounting on the rail

The fuel-pressure regulator (Figure 16) is of the diaphragm-controlled overflow type. A rubber-fabric diaphragm divides the pressure regulator into a fuel chamber and a spring chamber. Through a valve holder integrated in the diaphragm, the spring forces a movable valve plate against the valve seat so that the valve closes. As soon as the pressure applied to the diaphragm by the fuel exceeds the spring force, the valve opens again and permits just enough fuel to flow back to the fuel tank that equilibrium of forces is achieved again at the diaphragm.

The spring chamber is pneumatically connected with the intake manifold behind the throttle valve. As a result, the intake-manifold vacuum also has an effect in the spring chamber. There is therefore the same pressure ratio at the diaphragm as at the fuel injectors. This means that the pressure drop across the fuel injectors is solely a function of spring force and diaphragm surface area, and therefore remains constant.

Figure 16: Fuel-pressure regulator for mounting on the rail

1 Intake-manifold port (omitted for application in the fuel-supply module)
2 Compression spring,
3 Compression spring,
4 Valve holder,
5 Diaphragm,
6 Valve,
7 Nozzle,
8 Fuel supply,
9 Fuel return.

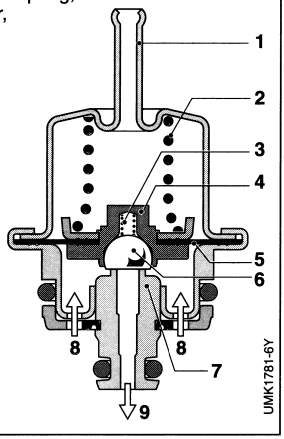
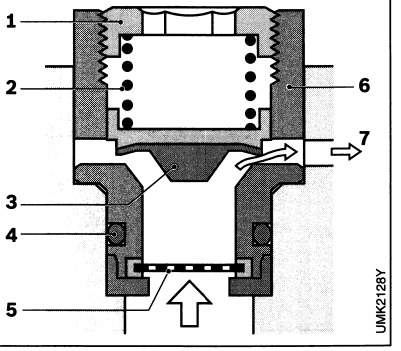


Figure 17: Fuel-pressure regulator for applications in the fuel-supply module
1 Adjusting screw (for primary pressure),
2 Spring,
3 Valve plate,
4 O-ring,
5 Prefilter,
6 Housing
7 Discharge to swirl pot.



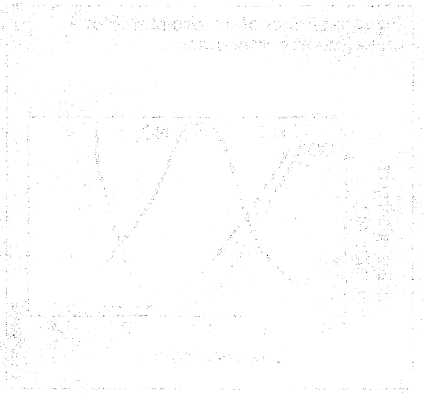
Installation in the fuel-supply module
For applications in the fuel-supply module the fuel-pressure regulator can have a simpler design that the pressure regulator mounted on the rail. These simple pressure regulators (Figure 17) have therefore become the accepted choice. The fuel flows through a coarse prefilter (protecting the pressure regulator against fouling) against the valve plate. An annular gap opens when the fuel pressure exceeds the spring force that presses the valve plate against the seal seat. The fuel flows through this annular gap and returns through the pressure regulator housing directly to the swirl pot.

Fuel-pressure attenuator

Function, design and operating principle
The repeated opening and closing of the fuel injectors, high-pressure pump pulsations (the principal cause today) and the periodic supply of fuel when electric positive-displacement fuel pumps are used lead to fuel-pressure oscillations. These can cause pressure resonances which adversely affect fuel-metering accuracy. It is even possible that under certain circumstances noise can be caused by these vibrations being transferred to the fuel tank and the vehicle bodywork through the mounting elements of the electric fuel pump, fuel lines, and fuel rail. These problems are alleviated by the use of special-design mounting elements and fuel-pressure attenuators.

The fuel-pressure attenuator has a similar design to the fuel-pressure regulator. Here too, a spring-loaded diaphragm separates the fuel chamber from the air chamber. The spring force is set in such a way that the membrane rises from its seat as soon as the fuel delivery pressure reaches its working range. This means that the fuel chamber is variable and not only absorbs fuel when pressure peaks occur, but also releases fuel when the pressure drops. In order to always operate in the most favorable range when the absolute fuel pressure fluctuates due to conditions at the manifold, the spring chamber can be provided with an intake-manifold connection.

Similar to the fuel-pressure regulator, the fuel-pressure attenuator can also be attached to the rail or installed in the fuel line. In the case of gasoline direct injection, it can also be attached to the high-pressure pump.



Mixture formation

Basic principles

A/F (Air/fuel) mixture

To be able to operate, a spark-ignition engine requires an air-fuel mixture. The mixture for ideal, theoretically complete combustion requires a mass ratio of 14.7:1 (stoichiometric ratio), i.e. 14.7 kg of air are required to burn 1 kg of fuel. Or: 1 *l* of fuel burns completely in roughly 9,500 *l* of air.

The specific fuel consumption is essentially dependent on the air-fuel ratio. Excess air is required in order to ensure genuine complete combustion, and thus as low a fuel consumption as possible. However, limits are imposed due to the flammability of the mixture and the available combustion time.

The A/F mixture also has a decisive impact on the efficiency of exhaust-gas treatment systems. State-of-the-art technology is represented by the three-way catalytic converter. However, it needs a specific stoichiometric A/F ratio in order to operate at maximum efficiency. This can reduce damaging exhaust gas components by higher than 99 %. The homogeneous engines available today are therefore operated with a stoichiometric mixture as soon as their operating status allows this. By way of departure, individual direct-injection engines are also operated with a lean mixture (stratified-charge operation) in the interest of fuel economy.

Certain engine operating statuses require mixture adaptation. Selective modifications to the mixture composition are required e.g. in the case of a cold engine.

Excess-air factor λ

The excess-air factor or air ratio λ (lambda) has been chosen to designate the extent to which the actual air/fuel mixture differs from the theoretically necessary mass ratio (stoichiometric ratio 14.7:1). λ denotes the ratio of air mass in the cylinder to the air requirement in the case of stoichiometric combustion.

Stoichiometric ratio

$\lambda = 1$: The inducted air mass corresponds to the theoretically required air mass.

Air deficiency

$\lambda < 1$: This indicates air deficiency and therefore a rich A/F mixture. The maximum engine output results at $\lambda = 0.85$ to 0.95.

Enriching the air-fuel mixture may be necessary to protect components against thermal overload (component protection). This reduces the exhaust-gas temperature, since the fuel is converted not completely to CO_2, but to CO. The enthalpy for the reaction to CO is lower than for complete conversion to CO_2. A lower thermal load e.g. for exhaust-gas turbocharger, λ sensor and catalytic converter results from the lower exhaust-gas temperature.

Figure 1: Effects of excess-air factor λ on specific fuel consumption b_e and uneven running at constant power output

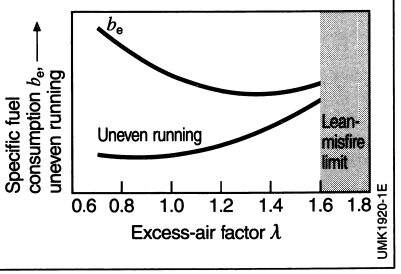
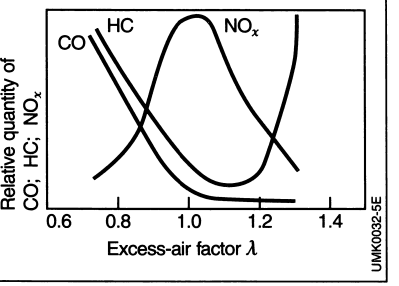


Figure 2: Effect of excess-air factor λ on pollutant composition



However, this also entails increased consumption. Enrichment is therefore only performed for high-load points for which an exhaust-gas temperature of over 900 °C would be reached.

Air surplus
$\lambda > 1$: In this range, there is an air surplus and thus a lean mixture.

This excess-air factor is characterized by reduced fuel consumption (Figure 1), but also by reduced power output. The maximum value for λ – the "lean-misfire limit" – that can be achieved is very strongly dependent on the construction of the engine and the mixture-formation system used. The mixture is no longer ignitable at the lean-misfire limit because of the lower number of HC molecules. Combustion misses occur, and this is accompanied by a marked increase in uneven running.

Influence of λ on fuel consumption and exhaust-gas composition
Spark-ignition engines achieve the lowest fuel consumption at constant engine output dependent on the engine at 20 to 50 % air surplus (λ = 1.2 to 1.5). This potential can however only be achieved with lean-burn concepts (homogeneous-lean or stratified-charge operation).

For a typical engine with intake-manifold injection, Figures 1 and 2 show the dependency of the specific fuel consumption and uneven running, as well as development of pollutants, on the excess-air factor at constant engine output. It can be deduced from these graphs that there is no ideal excess-air factor at which all the factors assume the most favorable value. In order to implement "optimal" consumption at "optimal" power output, excess-air factors of λ = 0.95 to 1.05 have proven conducive to the achievement of objectives for engines with intake-manifold injection and for homogeneous direct-injection engines – above all in view of exhaust-gas treatment with the three-way catalytic converter, which can optimally convert all three pollutant components only in a narrow range around $\lambda = 1$.

To achieve this, the inducted air mass is recorded precisely and an exactly metered fuel mass is apportioned by means of λ regulation in the engine-management system.

Intake-manifold injection
For optimum combustion in engines with intake-manifold injection, not only is a precise injected fuel quantity necessary, but also a homogeneous A/F mixture. This requires efficient fuel atomization. If this precondition is not satisfied, large fuel droplets will precipitate on the intake manifold or on the combustion-chamber walls. These large droplets cannot fully combust and will result in increased hydrocarbon emissions (HC).

Gasoline direct injection
Direct-injection engines with charge stratification involve different combustion conditions so that the lean-misfire limit occurs at significantly higher global λ values. In the part-load range, these engines can therefore be operated with considerably higher excess-air factors (up to λ = 4). Here too the area around the spark plug requires locally a stoichiometric mixture for good flammability.

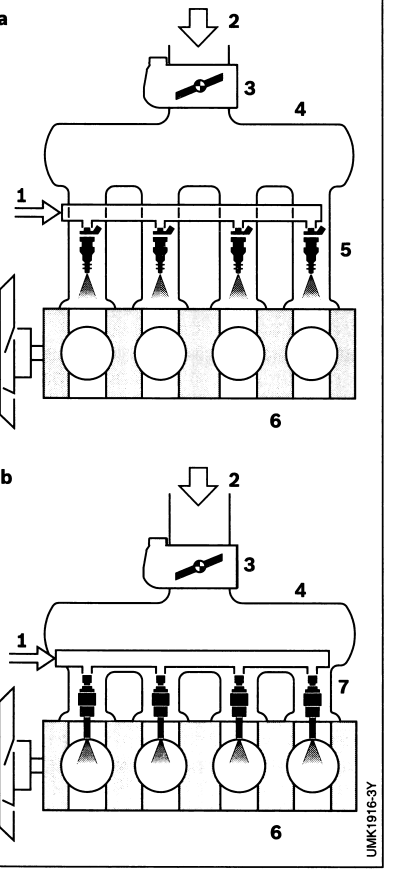
Unfortunately, in this case of a globally lean mixture, only two of the three ways of the three-way catalytic converter function. CO emissions and unburnt hydrocarbons can still be oxidized since sufficient oxygen is available. However, the third chief reaction, reducing the NO_x emissions, cannot be carried out sufficiently in lean-operation mode, since insufficient reduction partners are available. For this reason, in addition to catalytic exhaust-gas treatment with a three-way catalytic converter, a further exhaust-gas treatment facility must be installed for NO_x reduction (e.g. an NO_x storage catalytic converter).

Mixture-formation systems

It is the job of fuel-injection systems, or carburetors, to provide an A/F mixture which is adapted as well as possible to the relevant engine operating status. Without fuel-injection systems, particularly electronic systems, it would not be possible to maintain the increasingly narrowly defined limits for mixture composition and thus to adhere to the stipulations of the emission-control legislation. These requirements and the significant improvements in fuel metering and mixture formation together with the benefits in terms of fuel consumption, driving performance and power output have meant that, in the automotive sector, gasoline injection has completely superseded the carburetor in moder engines.

Until the beginning of this century, the automotive industry almost exclusively used systems in which A/F mixture formation takes place outside the combustion chamber (intake-manifold injection, Figure 3a). Systems with internal A/F mixture formation, i.e. where the fuel is injected directly into the combustion chamber (gasoline direct injection, Figure 3b), are designed to reduce fuel consumption even further and to increase power output, and have therefore become increasingly more important.

Figure 3: Schematic representation of fuel-injection systems
a) Intake-manifold injection,
b) Gasoline direct injection.
1 Fuel,
2 Air,
3 Throttle device,
4 Intake manifold,
5 Fuel injectors,
6 Engine,
7 High-pressure fuel injectors.



Intake-manifold injection

Gasoline injection systems for external mixture formation are characterized by the fact that the A/F mixture is created outside of the combustion chamber (in the intake manifold). The state of the art is represented by electronic intake-manifold injection systems where the fuel is injected cyclically for each individual cylinder directly ahead of the intake valves (Figure 4).

Systems that are based on mechanical-continuous fuel injection (K-Jetronic) or a central fuel injection arranged upstream of the throttle valve (Mono-Jetronic) are no longer of any significance for new developments.

Requirements

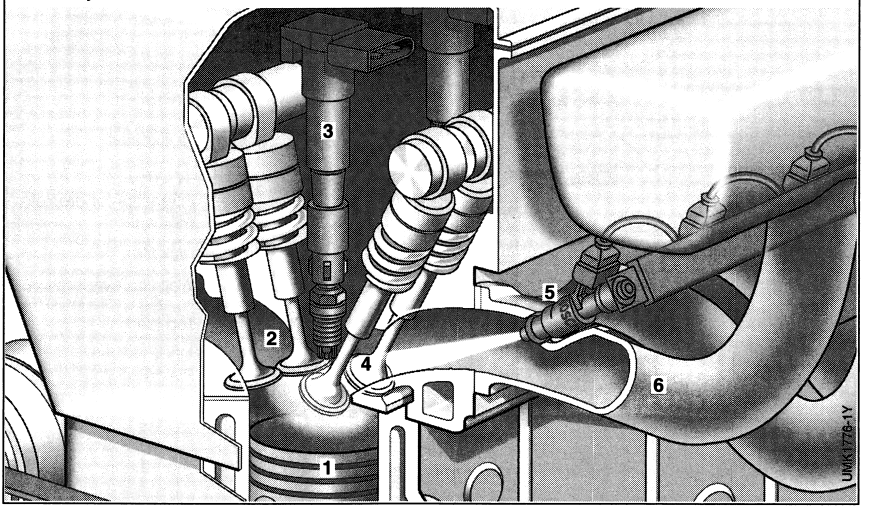
The high requirements for engine smooth-running and exhaust emissions make high demands on the A/F mixture composition in each power cycle. Precise injection timing is vital, as well as precise metering of the injected fuel mass as a factor of the engine intake air. In electronic multi-point fuel-injection systems, therefore, not only is each engine cylinder assigned an electromagnetic fuel injector, but this fuel injector is also activated individually for each cylinder. The engine control unit has the task of calculating both the required fuel mass for each cylinder and the correct start of injection for the fuel mass drawn in and the current engine operating status. The injection time required to inject the calculated fuel mass is a function of the opening cross-section of the fuel injector and the pressure differential between the intake manifold and the fuel supply system.

Fuel system

In intake-manifold injection systems, fuel is sent via the electric fuel pump (see Fuel supply), the fuel supply lines, and the fuel filter at typical system pressures of 3 to 7 bar to the fuel rail, which ensures that the fuel is evenly distributed to the fuel injectors. How the fuel is prepared by the fuel injectors is extremely important for the quality of the A/F mixture. It is essential that the fuel is atomized into very fine droplets. The spray shape and spray dispersal angle of the fuel injectors are adapted to the geometric shape of the intake manifold and cylinder head (see Intake-manifold fuel injector).

Figure 4: Principle of intake-manifold injection

1 Cylinder with piston, 2 Exhaust valves, 3 Ignition coil with spark plug, 4 Inlet valves, 5 Fuel injector, 6 Intake manifold.



Fuel injection

The exactly metered fuel mass is injected directly upstream of the cylinder intake valve(s). As the air enters through the throttle valve there is a pulse exchange with the finely atomized fuel, and the droplets on the surface mix increasingly with the air and therefore evaporate to a large extent (Figure 5). The required air-fuel mixture can therefore be formed at the right time by varying the point of injection.

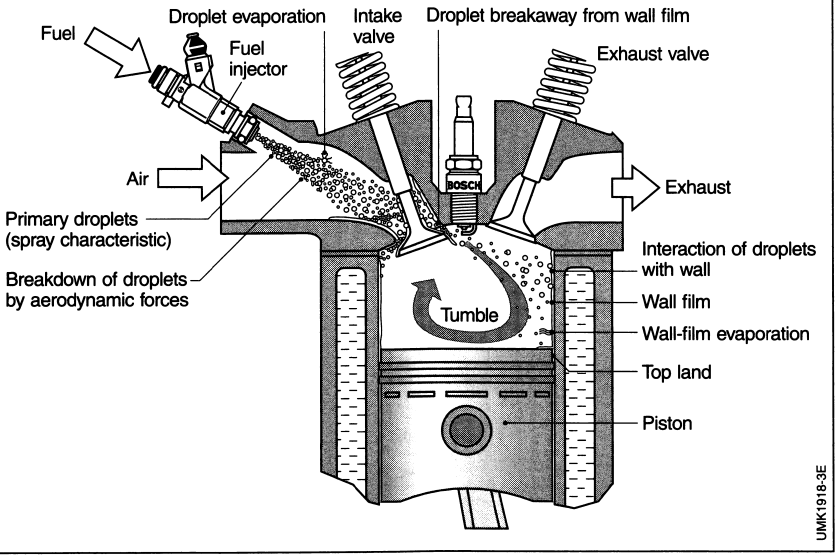
The amount of time available for mixture formation can be increased by injecting the fuel into the intake valves that are still closed. How much air-fuel mixture is created directly after injection depends on the quality of the fuel injector (primary droplet size). In this case, the primary evaporation occurs only when the intake valves open in response to the high flow velocities with an initially very small valve gap.

A proportion of the fuel precipitates on the intake duct (part of the intake manifold leading directly to the combustion chamber) and on the lower edge of the intake valves as wall film; the thickness of this film essentially depends on the pressure in the intake manifold, and thus on the engine load condition. In the case of non-stationary (dynamic) engine operation, this precipitation can lead to temporary deviations in the desired λ value ($\lambda = 1$), which means that the fuel mass stored in the wall film is to be kept as low as possible.

Wall-coating effects in the intake duct, particularly in cold-start conditions, cannot be neglected either. As fuel does not evaporate sufficiently, more fuel is required initially in the starting phase in order to create an ignitable A/F mixture. When the intake-manifold pressure then drops, parts of the previously formed wall film will vaporize. This may result in increased HC emissions if the catalytic converter is not running at operating temperature.

Irregular fuel injection may also result in the formation of wall films in the combustion chamber, and may in turn become critical emission sources. Defining the geometric alignment of fuel sprays ("spray targeting") will allow the selection of suit-

Figure 5: Mechanisms and factors influencing mixture formation in intake-manifold injection



able fuel injectors which will control and minimize manifold-wall fuel condensation in the area of the intake duct and the intake valves.

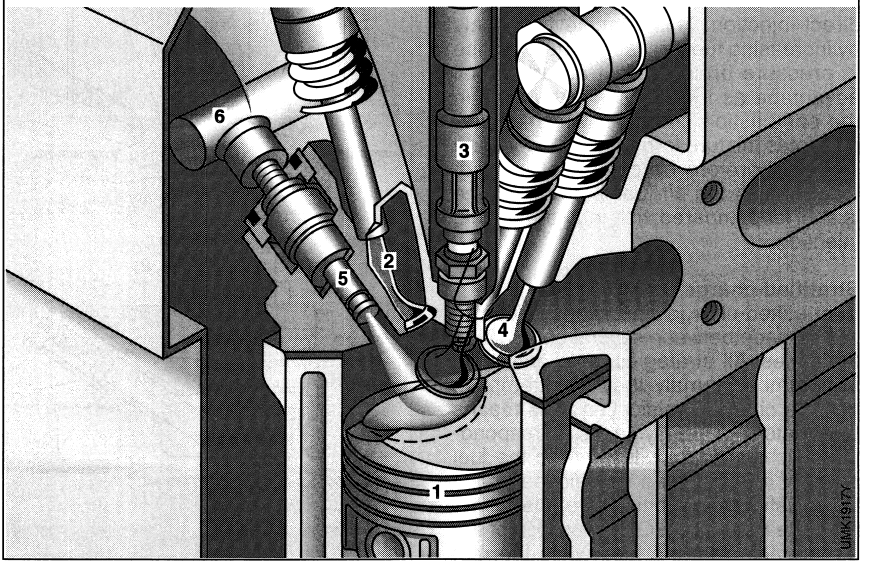
The intake manifolds used can be optimally adapted to the combustion air flow and the dynamic gas requirements of the engine.

Gasoline direct injection

In gasoline direct injection (GDI), unlike intake-manifold injection, pure air flows through the intake valves into the combustion chamber. Only then is the fuel injected into the air via an injector (see high-pressure fuel injector) located directly in the cylinder head (internal mixture formation, Figure 6). There are basically two main operating modes. Fuel injection in the intake stroke is called homogeneous operation, while fuel injection during compression is called stratified-charge operation. There are also various special modes, which are either a mixture of the two main operating modes or a slight variation of them.

In stratified-charge operation, the air is not restricted; the A/F mixture is lean. The excess air in the exhaust gas prevents conversion of the nitrogen oxides by means of a three-way catalytic converter. These direct-injection systems therefore require exhaust-gas treatment with an additional NO_x storage catalytic converter. For this reason, mostly direct-

Figure 6: Principle of gasoline direct injection
1 Cylinder with piston, 2 Intake valves, 3 Ignition coil with spark plug, 4 Exhaust valves, 5 High-pressure fuel injector, 6 Rail.



injection systems that work exclusively in the homogeneous mode are currently being placed on the market.

Homogeneous operation

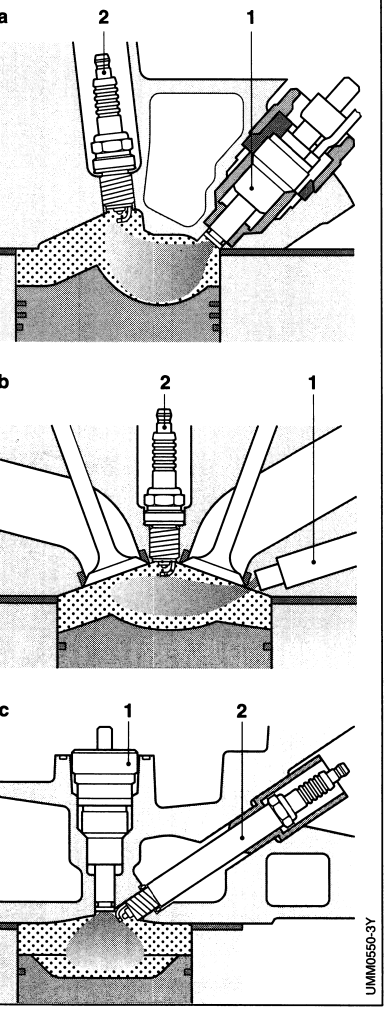
In homogeneous operation, mixture formation is similar to intake-manifold injection. The mixture is formed in a stoichiometric ratio ($\lambda = 1$). However, from a mixture formation point of view, there are some differences. For instance, there is no flow process around the intake valve to promote mixture formation, and there is much less time available for the mixture formation process itself. Whereas injection can take place across the overall 720° crankshaft (stored and synchronous with induction) of the four operating cycles in the case of intake-manifold injection, only an injection window of 180° crankshaft remains in the case of gasoline direct injection.

Fuel injection is only permitted in the induction stroke. This is because, prior to this, the exhaust valves are open and unburned fuel would otherwise escape into the exhaust-gas train. This would cause high HC emissions and catalytic converter problems. In order to deliver a sufficient quantity of fuel in this shortened period, the fuel flow through the injector must be increased for gasoline direct injection. This is achieved mainly by increasing fuel pressure. The increase in pressure brings with it an additional advantage as it increases turbulence in the combustion chamber, which in turn promotes mixture formation. The fuel and air can therefore be completely mixed, even though the air/fuel interaction time is shorter compared with intake-manifold injection.

Stratified-charge operation

In stratified-charge operation, a distinction is made between several combustion strategies. All strategies have one thing in common, namely they all attempt to achieve charge stratification. This means that, instead of supplying the corresponding stoichiometric air flow rate to the fuel quantity required for a specific load point by adjusting the throttle valve, the full air flow rate is supplied. Only a portion of the air interacts with the fuel, while the rest of the fresh air surrounds this stratified-charge cloud. This dethrottling action facilitates, together with the cooling effects of the directly injected fuel (evaporative emis-

Figure 7: Combustion processes for gasoline direct injection
a) Wall-directed combustion,
b) Air-directed combustion,
c) Spray-directed combustion.
1 High-pressure fuel injector,
2 Spark plug.



sion), an increase in compression and thus considerable fuel-saving potential.

Wall-directed combustion process
In a wall-directed combustion process, fuel is injected into the combustion chamber from the side (Figure 7a). A recess in the piston crown deflects the fuel spray in the direction of the spark plug. Mixture formation takes place on the path from the injector tip to the spark plug. As the mixture-formation time is very short during injection in the compression stroke (stratified-charge operation), the fuel pressure for this system must often be even higher than for homogeneous operation. The increased fuel pressure shortens the injection time and increases interaction with the air because pulse reflection is greater.

The disadvantage of the wall-directed combustion process is that fuel condenses on the piston wall, which increases HC emissions. As mixture-formation time is short, the stratified-charge cloud usually contains rich mixture zones at higher loads, and this increases the risk of soot production. At low loads, the pulse of the fuel droplets, which is used as a means of transporting the stratified-charge cloud to the spark plug, is low due to the small fuel mass. As a result, the flow must usually be restricted here so that the fuel meets with a lower density of air and thus a lower air resistance.

Air-directed combustion process
In principle, an air-directed combustion process works in exactly the same way as a wall-directed process. The main difference is that the stratified-charge cloud does not interact directly with the piston recess. Instead, it moves on a cushion of air (Figure 7b). The cushion of air is generated by the air present in the cylinder. In contrast to the wall-directed combustion process, with air-directed combustion the fuel-spray injection angle is much flatter such that the fuel spray cannot penetrate the full amount of air up to the piston. This solves the problem of fuel condensing on the piston recess.

Compared with wall-directed combustion processes, the air flows in air-directed combustion processes cannot be fully reproduced. This gives rise to large variations from injection to injection and thus frequently to poor combustion stability right through to individual misfires.

Often, the actual combustion processes are a mixture of wall-directed and air-directed processes, depending on the operating point in each case.

Spray-directed combustion process
The spray-directed combustion process is visually different from the other two processes in that the injector is installed at a different location. It is located top center and injects vertically down into the combustion chamber (Figure 7c). The fuel spray is not deflected. As a result of the high injection pressure (200 bar), an air-fuel mixture forms around the fuel spray already during the injection process. Located directly next to the injector is the spark plug, which ignites the air-fuel mixture.

As a result, with the spray-directed combustion process the mixture-formation time is very short. This requires an even higher fuel pressure for the spray-directed combustion process. This combustion process can eliminate the disadvantages of fuel condensing on the manifold walls, air-flow dependency, and flow restriction at low loads. It therefore has the greatest potential for fuel saving. Nevertheless, the short mixture formation time is a huge challenge for the fuel-injection and ignition systems.

Other operating modes
In addition to homogeneous and stratified-charge operation, there are also special operating modes. They include "changeover between operating modes" (homogeneous-stratified mode), "catalytic-converter heating", and "knock protection" (homogeneous-split mode), and "homogeneous-lean mode". These operating modes are sometimes very complex and only ever represent short, temporary states. Therefore they are frequently not operating modes for longer operation.

Fuel injectors

Essentially, the mixture-formation components must ensure that the air-fuel mixture is formed in the most optimal way possible to facilitate combustion. In intake-manifold injection, this is mainly the task of the fuel injector, while in gasoline direct injection, the high-pressure fuel injector can be assisted by a turbulence flap.

Essentially, there are two different ways of actuating the fuel injectors. If one injector is operated by means of a conventional electromagnet, the other has a piezoelectric stack serving as the actuator.

Intake-manifold fuel injector

Design and function

Essentially, electromagnetic fuel injectors are comprised of the following components: the valve housing with electrical connection and hydraulic port, the solenoid coil, the moving valve needle with solenoid armature and valve ball, the valve seat with injection-orifice plate, and the valve spring.

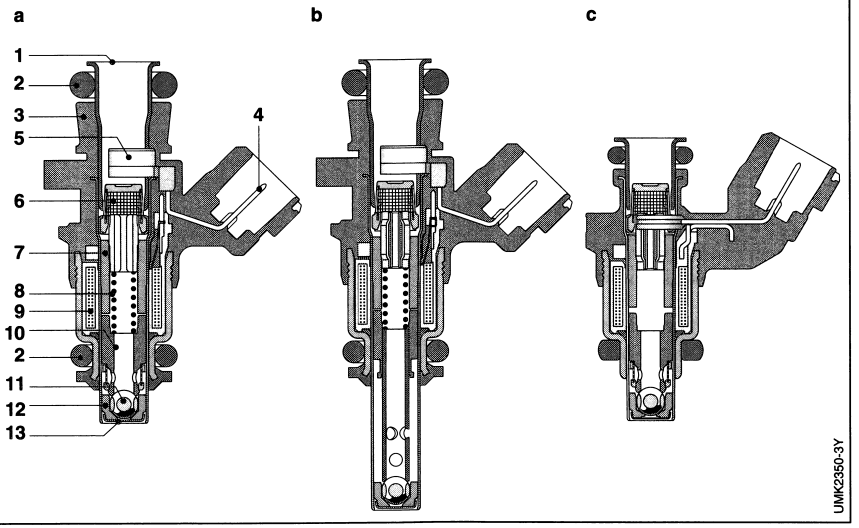
Figure 8a and Figure 8b show the EV14 fuel injector, as is used in intake-manifold injection. The new reduced-size fuel-injector variant EV-SE (Figure 8c) has been developed for small installation spaces and small engines (e.g. for two-wheeled vehicles). Both injector variants are also available with an extended injection point.

To adapt the installation position optimally to the intake-manifold and intake-port geometries, different EV14 design variants are available to choose from: Three overall lengths are available with different electrical connector variants and optionally with an extended injection point (Figure 9.)

A filter strainer in the fuel inlet protects the fuel injector against contamination. The sealing ring (O-ring) on the hydraulic port seals off the injector at the fuel rail. The lower sealing ring provides the seal between the injector and the intake manifold.

Figure 8: Fuel injectors (Bosch) for intake-manifold injection

a) Fuel injector, standard version EV-14 fT (flat tip),
b) Fuel injector EV14-xT (extended tip) with extended injection point,
c) Fuel injector EV-SE of reduced size for use e.g. in motorcycles.
1 Hydraulic port, 2 Sealing ring (O-ring), 3 Valve housing, 4 Electrical connection, 5 Connector clip, 6 Filter strainer, 7 Internal pole (stop for solenoid armature), 8 Valve spring, 9 Solenoid coil, 10 Valve needle with solenoid armature, 11 Valve ball, 12 Valve seat, 13 Injection-orifice plate.



When the solenoid coil is de-energized, the valve needle and valve ball are pressed against the cone-shaped valve seat by the valve spring and the force exerted by the fuel pressure. The fuel-supply system is thus sealed off from the intake manifold.

When the solenoid coil is energized, this generates a magnetic field which attracts valve-needle solenoid armature. The valve ball lifts off the valve seat and the fuel is injected. When the excitation current is switched off, the valve needle closes again due to spring force.

The fuel is atomized by an injection-orifice plate. The orifices are stamped out of the plate and ensure that the injected fuel quantity remains highly reproducible. The spray pattern and the atomization quality of the fuel leaving the injector is produced by the number of orifices and their configuration (maximum twelve).

The injected fuel quantity per unit of time is determined mainly by the system pressure in the fuel-supply system (typically 3 to 4 bar, corresponding to 300 to 400 kPa), the back pressure in the intake manifold, and the geometry of the fuel-exit area.

Spray formation and spray targeting

An injector's spray formation, i.e. its spray shape, spray angle, and fuel-droplet size, influences the formation of the air/fuel mixture. Individual geometries of intake manifold and cylinder head make it necessary to have different types of spray formation. Different spray-formation variants are available in order to satisfy with these requirements.

Tapered spray

Individual fuel sprays emerge through the openings in the injection-orifice plate (Figure 10a). These fuel sprays combine to form a spray taper.

Tapered-spray injectors are typically used in engines which feature only one intake valve per cylinder.

Dual spray

Dual-spray formation (Figure 10b) is as a rule used in engines which feature two intake valves per cylinder. The holes in the injection-orifice plate are arranged in such a way that two fuel sprays – which are each formed from a number of individual sprays that produce a homogeneous spray shortly after leaving the injection or-

Figure 9: Three overall lengths for the EV14 fuel injector, optionally with extended injection point

a) EV14, fuel injector, compact, extended injection point,
b) EV14, fuel injector, compact,
c) EV14, fuel injector, standard,
d) EV14, fuel injector, long.

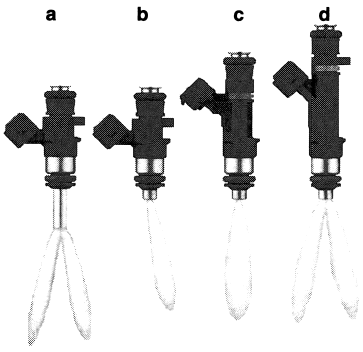
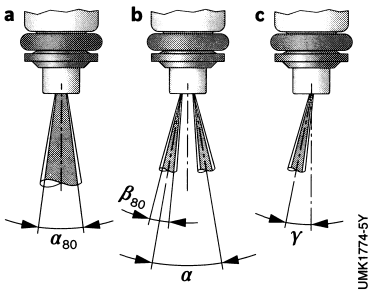


Figure 10: Spray geometries

a) Tapered spray,
b) Dual spray,
c) Gamma angle.
α_{80} 80 % of fuel is inside the angle α,
α Angle between two individual sprays,
β_{80} 80 % of fuel is inside the angle β,
γ Offset spray angle.

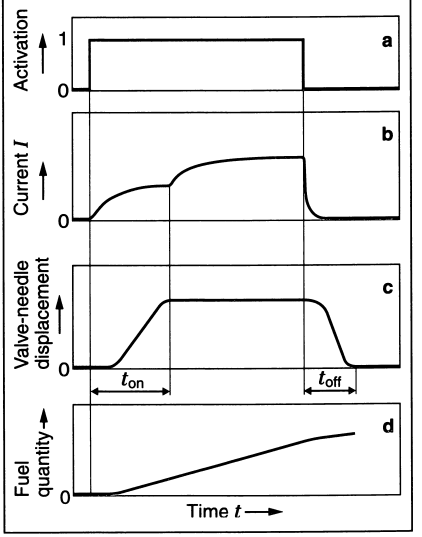


ifices – impact ahead of the intake valves or against the web between the intake valves.

Gamma angle
Compared to the injector's principle axis, the fuel spray in this case (single spray and dual spray) is at an angle: the offset spray angle (Figure 10c).

Injectors with this spray shape are mostly used when installation conditions are difficult, for example when the injectors are installed so steeply in the intake manifold that the required spray targeting can only be achieved by the use of a gamma-angle spray.

Figure 11: Activation of fuel injector
a) Activation signal,
b) Current curve,
c) Valve-needle lift,
d) Injected fuel quantity.
t_{on} Pickup time,
t_{off} Dropout time.



Electrical activation
The output module in the engine ECU actuates the injector with a switching signal (Figure 11a). The current in the solenoid coil rises (Figure 11b) and causes the valve needle to lift (Figure 11c). After the pickup time has elapsed, the maximum valve-needle lift is achieved at the solenoid-armature stop at the internal pole. Depending on the solenoid-valve design, the current continues to rise after the armature stop because the magnetic circuit at this point is not yet at 100 % saturation. The kink in the current curve is triggered by the armature stop.

Fuel is sprayed as soon as the valve ball lifts off its seat. The quantity of fuel injected during an injection pulse is shown in Figure 11d.

Because the magnetic field is not decreased abruptly after activation is shut down, the valve closes with a delay. The valve is fully closed again after the dropout time has elapsed.

The non-linearity during the valve pickup and dropout phases must be compensated for throughout the period that the injector is activated (injection duration). The speed at which the valve needle lifts off its seat is also dependent on the battery voltage. Battery-voltage-dependent injection-duration extension corrects these influences.

Electromagnetic high-pressure fuel injector

The high fuel pressure required for gasoline direct injection (nominal pressure up to approximately 35 MPa) and the installation location give rise to extended requirements for the fuel-injector component when compared with intake-manifold fuel injectors.

Design and function
The function of the high-pressure fuel injector is to meter and atomize the fuel. Atomization ensures that the fuel is quickly mixed with the air in the combustion chamber. In the process, the A/F mixture is formed within a physically restricted area.

The high-pressure fuel injector (Figure 12) consists of the individual components:

- housing with hydraulic port and electrical connection, tolerance compensation and acoustic decoupling element, and a locking element,
- valve seat with injection orifices and pre-stages,
- valve needle with mechanically isolated solenoid armature and spring,
- return spring with setting sleeve,
- coil with conductor bar.

Depending on the installation position in the cylinder head – lateral or central – either short or long valves are used. The long valve is typically based on a short valve which is adapted to the required length with an extension tube for the hydraulic port and an extended conductor bar for the electrical connector. The current generation has an valve-tip diameter reduced to 6 mm to satisfy the requirements with regard to installation space in downsizing engines.

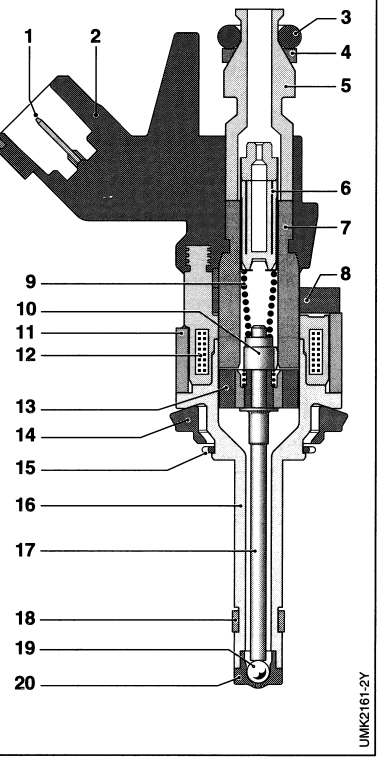
When current flows through the coil, it generates a magnetic field in conjunction with the other components of the magnetic circuit. This magnetic field accelerates the solenoid armature when the valve needle is still at rest. When the solenoid armature contacts the upper stop, the latter lifts the valve needle against spring force and the hydraulic load off the valve seat and opens the valve outlet opening. As the fuel pressure is considerably higher than the combustion-chamber pressure, the fuel is injected into the combustion chamber.

When the current is switched off, the valve needle and the hydraulic load on the valve needle are pushed back into the seat by the return spring, and this stops the injection process. The solenoid armature is returned by the spring to its rest position on the lower stop.

Reproducible fuel quantities can be metered by defining the opening and keeping the opening cross-section constant with the needle fully raised. Here, the fuel quantity is dependent on the pressure in the fuel rail, the backpressure reaction in the combustion chamber, and the valve opening time.

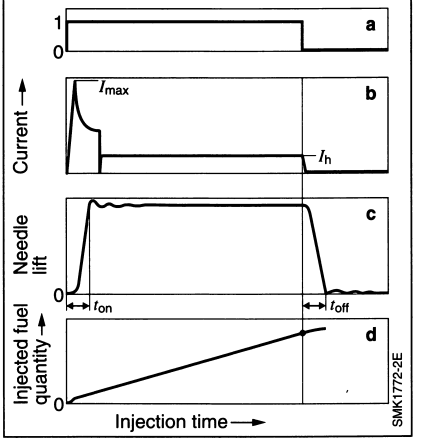
Figure 12: Electromagnetic high-pressure fuel injector (Bosch) for gasoline direct injection

1 Electrical connection,
2 Extrusion coating,
3 O-ring, 4 Support plate,
5 Conductor barrel,
6 Filter with setting sleeve,
7 Internal pole,
8 Cover,
9 Compression spring,
10 Stop ring,
11 Magnetic sleeve,
12 Solenoid coil,
13 Solenoid armature,
14 Support element,
15 Snap ring,
16 Valve housing,
17 Valve needle,
18 Sealing ring,
19 Valve ball,
20 Valve seat with injection orifices.



Spray formation
Predominantly multi-hole valves are used to prepare the fuel. By suitable spray configuration the injection-orifice alignment is adapted to match the installation position to the respective combustion chamber. Here the sprays must be configured in such a way that fuel and fresh air mix thoroughly in the combustion chamber and in the process valves, pistons and combustion-chamber walls are moistened with fuel as little as possible. The fuel injectors are typically configured with five to seven sprays, depending on the desired throughflow, which is configured to match the output per cylinder and thereby determines the mean orifice diameter. The penetration depth of the sprays can be reduced for individual injection orifices by a diameter that is smaller than the mean diameter of all the injection orifices. This makes it possible to reduce the moistening of the piston and the combustion-chamber walls with liquid fuel. Other sprays with a larger orifice diameter are provided to achieve the required throughflow. The injection orifice is for the most part supplemented by a pre-stage. The purpose of this pre-stage is to keep the injection orifice free from deposits which are produced depending on the boundary operating and environmental conditions and on the fuel used.

Figure 13: Signal curves for activating the high-pressure fuel injector
a) Triggering signal calculated by the ECU,
b) Current curve in the high-pressure fuel injector,
c) Valve-needle lift,
d) Injected fuel quantity.
I_{max} Maximum current in the booster phase,
I_h Holding current,
t_{on} Pickup time,
t_{off} Dropout time.



Electrical activation
To ensure a defined and reproducible injection process, the high-pressure fuel injector is activated by a complex current curve. The engine control unit delivers a digital signal for this purpose. An output module uses this signal (Figure 13a) to generate the actuating signal (Figure 13b) for the high-pressure fuel injector.

A DC/DC converter in the engine ECU generates the booster voltage of 65 V. This voltage creates in conjunction with the booster capacitor a high current at the beginning of the activation process, thus ensuring that the solenoid armature and the valve needle are raised quickly. In the pickup phase, the valve needle then achieves the maximum opening lift (Figure 13c). When the fuel injector is open (maximum valve-needle lift), a lower control current (holding current) is sufficient to keep the valve open. With a constant injector-needle lift, the injected fuel quantity is proportional to the injection duration (Figure 13d).

The CVO process (Controlled Valve Operation) has been specially developed to reduce the quantity scatter from valve to valve. The opening and closing times of the fuel injector can be determined from the electrical variables with this mechatronic approach. The current actual open duration is formed from these measured variables; this is delivered to a controller, where it is compared with a setpoint open duration. In this way, it is possible particularly for very small quantities to achieve high metering precision and maintain this level of precision with a downstream adaptation over the entire service life.

During multiple fuel injection the high-pressure fuel injector is actuated more than once during an engine cycle, reducing the level to which the internal en-

gine walls are moistened with liquid fuel. For the fuel injector this means achieving short idle times and more numbers of load cycles to be withstood over the operating time compared with single injection.

Piezo high-pressure fuel injector

Area of application

The piezo high-pressure fuel injector has been used since 2005 in complex combustion processes which impose specific requirements on the quality of the fuel spray and its stability over the engine's service life. Moreover, it combines outstanding very-small-quantity metering precision with high maximum injected fuel quantities and thus makes possible high specific power output without having to make compromises at low-load operating points. An example of highly-developed combustion processes is the spray-guided stratified-lean combustion process, which is achieved exclusively with the piezo high-pressure fuel injector in production vehicles.

Design and function

The piezo high-pressure fuel injector is operated at 20 MPa (200 bar) fuel pressure and comprises the function elements of valve assembly, actuator module, compensation element ("coupler"), and housing and connecting parts (Figure 14).

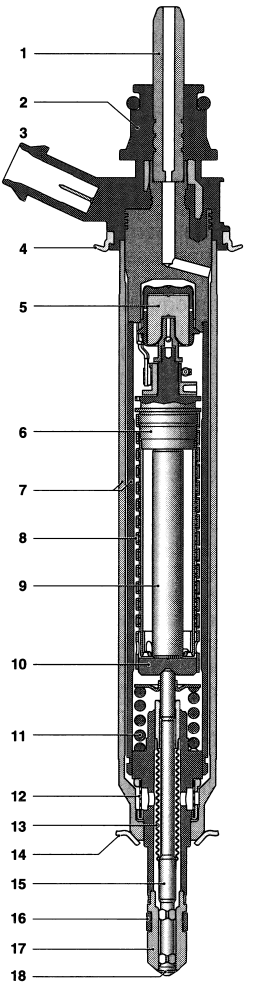
The valve assembly admits injected fuel into the combustion chamber and defines the shape of the spray. It essentially consists of the outward-opening valve needle which is pretensioned by a closing spring against a housing body. The nominal valve-needle lift (full lift) is 33 µm and is achieved after 180 µs. The closing operation also takes 180 µs from full lift.

Further important components of the valve assembly are a fine filter in the inlet to protect the nozzle against particulates and bellows to isolate the fuel from the dry actuator chamber.

Installed in the actuator chamber is the piezo actuator, which through its change of length under voltage opens and closes the valve needle and thus substantially determines fuel-quantity metering. It consists of over 400 active layers and is pretensioned by a tube spring. Aside from the actuator and tube spring, the electrical

Figure 14: Function elements of piezo fuel injector (Bosch)

1 Hydraulic port, 2 Plastic coating,
3 Electrical connection, 4 Injector-duct seal,
5 Linear compensation element,
6 Actuator foot, 7 Housing,
8 Piezo-actuator pretensioning spring,
9 Piezo actuator, 10 Actuator head,
11 Needle closing spring, 12 Mesh filter,
13 Metal bellows, 14 Support element,
15 Valve needle, 16 Combustion-chamber seal,
17 Valve body, 18 Injection nozzle.



contacting and insulation of the actuator against the electrical components are the essential elements of the actuator module. The direct non-positive connection between actuator module and valve needle provides for very short activation times from 75 µs. In the event of such very short activations the valve needle does not achieve its full lift.

The coupler exerts in the switching chain (comprising the valve needle, the actuator module, and the coupler itself) a defined force which counteracts the closing-spring force. In the coupler two steel diaphragms hold a hydraulic fluid which moves a working piston by means of a slow movement between two cavities. This brings about a compensation of the different thermal expansions of the metallic and ceramic components with the result that the coupler force within the switching chain remains to a large extent constant over all the thermal operating conditions of the high-pressure fuel injector. In this way, a reproducible valve-needle lift and thus a reproducible injected fuel quantity in the different engine operating states are made possible. Fast, short changes of switching-chain length by activation of the piezo actuator are not compensated by the coupler in contrast to changes of switching-chain length as a result of slow thermal effects. It works in this respect like a high-pass filter.

The above-mentioned modules are mounted together with the housing components including the fuel supply. A coding process integrated in the production sequence determines the individual charge demand of each piezo high-pressure fuel injector in order to compensate production tolerances. Compensating the production tolerances improves the precision of the fuel-quantity metering. The charge demand is written in the form of a data matrix code to the fuel injector and read later into the engine ECU.

Spray formation and spray targeting

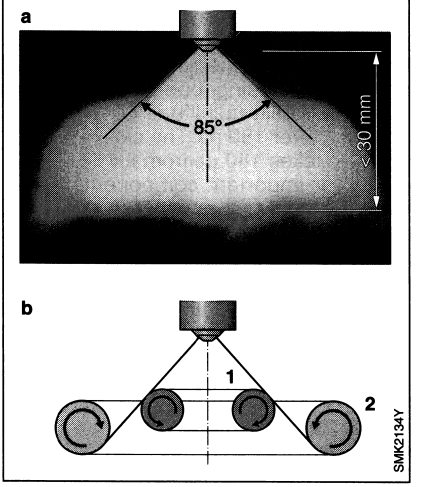
The spray of an outward-opening nozzle differs by virtue of the principle involved fundamentally from that of a multi-hole nozzle: The annular orifice created when the valve needle opens causes a tapered discharge of spray with droplet sizes in the region of 8 to 13 µm. A rotationally symmetric hollow taper with a nominal spray angle of 86 ° is created. Spatially permanently assigned boundary vortices, which can be used in the course of a combustion process, form both outside and inside along the taper surfaces (Figure 15). The spray shape is stable over a large combustion-chamber-pressure and charge-movement range such that even injections that shortly follow each other are possible in the compression stroke up to shortly before ignition (thus at high combustion-chamber pressures). The spray is, thanks to the design of the nozzle, very robust against disturbance, deposit, and aging influences. The is advantageous particularly for the stratified combustion process mentioned at the start, but also for example in catalyst-heating mode.

Electrical activation

The piezo high-pressure fuel injector is operated with up to 200 V. Activation is performed by an output stage integrated in the ECU, with the voltage generated by a

Figure 15: Tapered spray pattern of piezo fuel injector

Fuel pressure 200 bar,
chamber pressure 6 bar (lab measurement).
a) Realistic representation,
b) Schematic representation.
1 Inner boundary vortex,
2 Outer boundary vortex.



DC/DC converter (Figure 16). Activation is charge-controlled: The injector is opened by a specifically targeted charging operation (nominal value for full lift 0.69 mC) and closed by an analogous discharging operation. The activation duration (75 to 5000 µs) serves as the primary activation variable, the charge value to be achieved and the charging and discharging speeds are parameterizable. Varying the pressure to obtain an injected-fuel-quantity spread between minimum and maximum quantity is, due to the possible extremely short activation duration and on account of the high hydraulic cross-section at full lift, not necessary: Injected fuel quantities between 0.5 mg and 150 mg can be covered by a fuel injector per working cycle. The injection characteristic is linear right through to the smallest injected fuel quantities (Figure 17). Idle times between two opening operations right down to 50 µs can be achieved; this is considerably shorter than with a fuel injector with a solenoid-operated mechanism.

Finally, a number of adaptation functions which measure the fuel injector in the installed state are used in engine operation to guarantee very-small-quantity metering precision of the injector's service life. As well as model-based compensations of the temperature-dependent reactions of the actuator to an electric charge, it is also possible to compensate on a cylinder-specific basis very small quantity deviations over the injector's service life which if uncompensated would generate torque differences. The compensations act on the basis of the causative principle either on the valve-needle lift or on the activation duration.

Figure 16: Activation of piezo fuel injector
a) For full lift,
b) For partial lift.

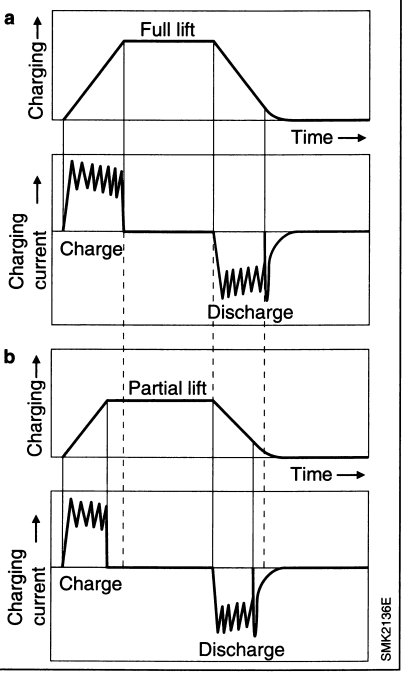
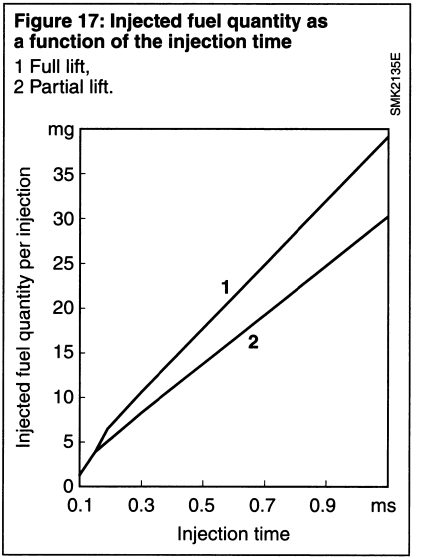


Figure 17: Injected fuel quantity as a function of the injection time
1 Full lift,
2 Partial lift.



Ignition

Basic principles

Function

On a spark-ignition (SI) engine, the combustion process is initiated by an externally supplied ignition. The ignition is responsible for igniting the compressed A/F mixture at the right time. This is done by producing an electric spark between the electrodes of a spark plug in the combustion chamber.

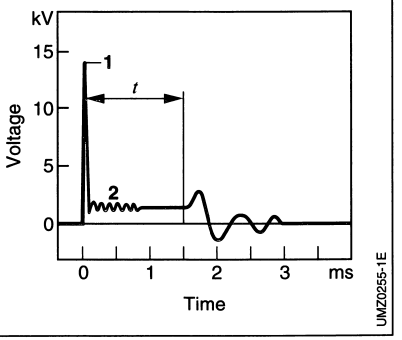
Consistent, reliable ignition under all conditions is essential to ensure fault-free engine operation. Misfiring leads to combustion misfires and damage or destruction of the catalytic converter, poor exhaust emission figures, higher consumption and lower engine performance.

Ignition spark

An electric spark can only occur at the spark plug if the necessary ignition voltage is exceeded (Figure 1). The ignition voltage is dependent on the spark-plug electrode gap and the density of the A/F mixture at the moment of ignition. It can be up to several 10 kV. After flashover, the voltage at the spark plug drops to the firing voltage. The firing voltage depends on the length of the spark plasma (electrode gap and excursion by the A/F mixture flow) and can rise from several 100 V to a few kV.

During the ignition-spark combustion time (spark duration), the ignition-system energy is converted into the ignition spark. After the spark has broken away, the voltage is damped and drops to zero.

Figure 1: Spark-plug voltage characteristic with static or semi-static A/F mixture
1 Ignition voltage,
2 Firing voltage,
t Spark duration.



Mixture ignition and ignition energy

The electrical spark between the spark-plug electrodes generates a high-temperature plasma. If mixture conditions at the spark plug are suitable and sufficient energy has been supplied by the ignition system, the resulting arc develops into a flame front that propagates independently.

The ignition must guarantee this process under all engine operating conditions. Under ideal conditions (e.g. in a simulation of a combustion chamber, a "combustion bomb"), provided that the A/F mixture is stationary, homogeneous and stoichiometric, an energy of approx. 0.2 mJ is required to ignite the mixture by means of electric spark for each individual ignition process. In real engine operation, however, much higher energy levels are required. Some of the spark energy is converted at flashover, and the rest in the spark combustion phase.

Larger electrode gaps generate a larger arc, but require higher ignition voltages. Lean A/F mixtures or turbocharged engines have a higher ignition-voltage demand. At a given level of energy, the spark duration shortens as the ignition voltage increases. A longer spark duration generally stabilizes the combustion; lack of mixture homogeneity at the moment of ignition in the area of the spark plug can be compensated for by means of a longer spark duration. A/F-mixture turbulences such as those which occur in the stratified-charge mode with gasoline direct injection can divert the ignition spark to such an extent that it extinguishes. Follow-up sparks are then required to ignite the A/F mixture again.

The need for higher ignition voltages, longer spark durations, and the provision

of follow-up sparks have resulted in the design of ignition systems with higher ignition energy. If not enough ignition energy is produced, inflammation will not occur, resulting in engine misfire. The system must therefore deliver enough ignition energy to ensure reliable ignition of the A/F mixture under all operating conditions.

Efficient fuel atomization and good access of the A/F mixture to the ignition sparks enhance ignitability, extend spark duration and spark length, and lengthen the electrode gap. The dimensions of the spark plug determine the position (see Spark position, spark plug) and length of the spark; spark duration depends on the type and design of the ignition system, as well as on the instantaneous ignition conditions in the combustion chamber. Dependent on the engine requirements (intake-manifold injection, gasoline direct injection or exhaust-gas turbocharging), the spark energy of ignition systems lies in a range of approx. 30 to 100 mJ.

Moment of ignition

The start of combustion in a spark-ignition (SI) engine can be controlled by selecting the moment of ignition. The moment of ignition always refers to the top dead center of the spark-ignition (SI) engine power cycle. The earliest possible moment of ignition is determined by the knock limit, and the latest possible moment of ignition by the combustion limit or the maximum permissible exhaust-gas temperature. The moment of ignition affects the torque output, exhaust-gas emissions and fuel consumption.

Basic ignition point

Maximum combustion, and thus maximum combustion pressure, should occur shortly after top dead center (Figure 2) to deliver maximum engine torque. Since a certain amount of time passes from the moment of ignition to complete combustion of the A/F mixture, ignition must occur before top dead center. The moment of ignition must be advanced as rotational speed increases and charge decreases.

Likewise, the moment of ignition must be advanced in the case of lean A/F mixtures ($\lambda > 1$) because the flame front propagates more slowly. Recirculated or retained exhaust gas (external or internal exhaust-gas recirculation) has a compa-

Figure 2: Combustion-chamber pressure characteristic for various moments of ignition

1 Ignition (Z_a) at the correct point in time,
2 Ignition (Z_b) too advanced,
3 Ignition (Z_c) too retarded.

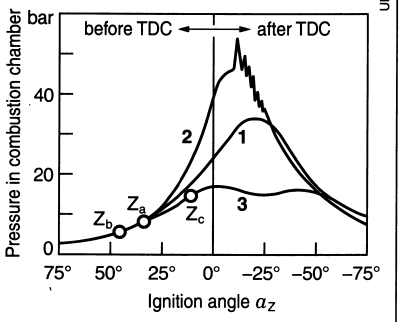
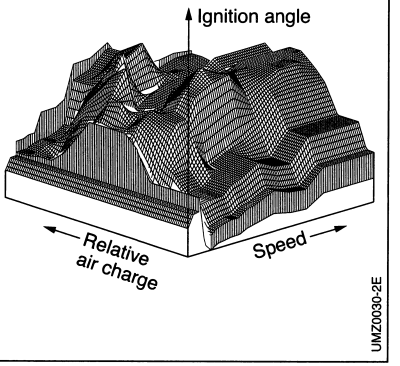


Figure 3: Ignition map as a function of engine speed and relative air charge



rable effect. Ignition-timing adjustment, therefore, essentially depends on rotational speed, charge, and the air/fuel ratio (excess-air factor λ) or in today's customary stoichiometric homogeneous engines the amount of residual exhaust gas in the cylinder. The moments of ignition are determined on the engine test bench and, in the case of electronic engine management systems, are stored in program maps (Figure 3).

Ignition-timing corrections and operating-point-dependent moments of ignition

Electronic engine-management systems can take other effects on the moment of ignition into consideration in addition to rotational speed and charge. The basic moment of ignition can either be modified by means of additive corrections or replaced for certain operating points or ranges by special ignition angles or ignition maps. Examples of ignition-timing corrections are knock control, the correction angle for the gasoline direct injection homogeneous-lean operation, and warm-up. Examples of special ignition angles or ignition maps are gasoline direct injection stratified-charge operation, and starting operation. Final implementation depends on the prevailing application.

Exhaust gas and fuel consumption

The moment of ignition has a considerable impact on exhaust gas because it can be used to control the various untreated exhaust-gas constituents directly. However, the various optimization criteria, such as exhaust gas, fuel economy, drivability, etc., may not always be compatible, so it is not always possible to derive the ideal moment of ignition from them.

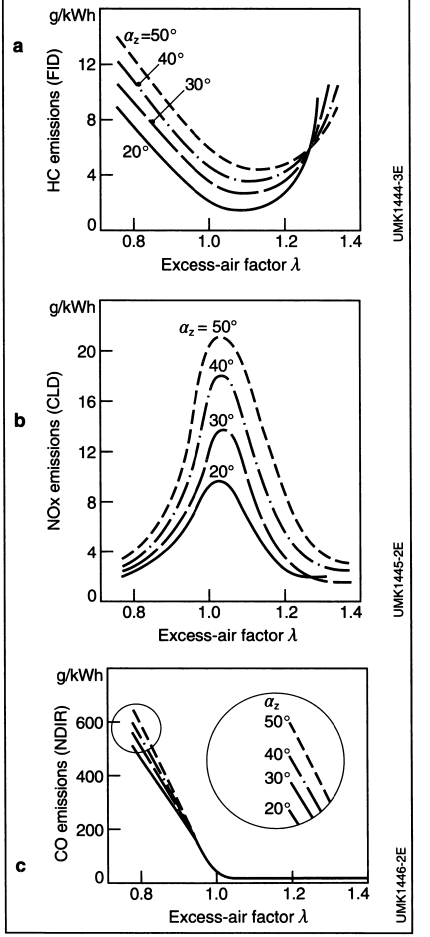
Shifts in the moment of ignition induce mutually inverse response patterns in fuel consumption and exhaust-gas emissions (Figure 4 and Figure 5). Whereas more spark advance increases torque and therefore power and reduces fuel consumption, it also raises HC and, in particular, nitrogen-oxide emissions. Excessive spark advance can cause engine knock that may damage the engine. Retarded ignition results in higher exhaust-gas temperatures, which can also harm the engine.

Electronic engine-management systems featuring programmed ignition curves are designed to adapt the moment of ignition in response to variations in factors such as rotational speed, load, temperature,

Figure 4: Influence of excess-air factor λ and moment of ignition α_z on pollutant emissions

a) Hydrocarbon emissions (HC),
b) Nitrogen-oxide emissions (NO_x),
c) Carbon-monoxide emissions (CO).

Flame Ionization Detector (FID)
Chemiluminescence Detector (CLD)
Non-Dispersive Infrared Detector (NDIR)



etc. They can thus be employed to achieve the optimum compromise between these mutually antagonistic objectives.

Knock control

Basic principles

Electronic control of the moment of ignition offers the possibility of accurate control of the ignition angle as a function of rotational speed, load, temperature, etc. Nevertheless, if there is no knock control, there must still be some means to define a clear safety margin to the knock limit.

This margin is necessary to ensure that, even in the most knock-sensitive case with regard to engine tolerances, engine aging, environmental conditions, and fuel quality, no cylinder can reach or exceed the knock limit. The resulting engine design leads to lower compression, retarded moments of ignition, and thus worsening fuel consumption and torque.

These disadvantages can be avoided through the use of knock control. Experience shows that knock control increases engine compression. This results in lower fuel consumption and higher torque. Now, however, the pilot control ignition angle no longer has to be determined for the conditions most sensitive to knocking but rather for the conditions least sensitive to knocking (e.g. compression of the engine at lowest tolerance limit, best possible fuel quality, cylinder least sensitive to knocking). Each individual engine cylinder can now be operated throughout its service life in virtually all operating ranges at its knock limit, and thus at optimum efficiency. For this type of ignition angle adjustment, a reliable method of knock detection is essential. It should detect knock for each cylinder throughout the engine's operating range starting from a specified knock intensity.

Knock-control system

A knock-control system consists of knock sensor, signal evaluation, knock detection, and ignition-angle control system with adaptation facility (Figure 6).

Figure 5: Influence of excess-air factor λ and moment of ignition α_z on fuel consumption and torque
a) Torque,
b) Specific fuel consumption.

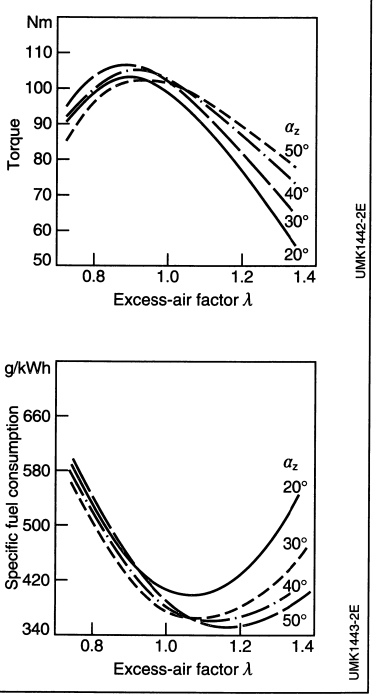
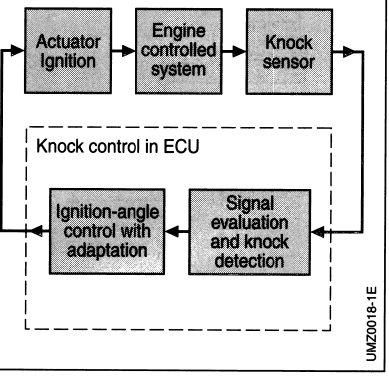


Figure 6: Schematic of a knock-control system



Knock sensor
A typical symptom of combustion knock is high-frequency vibrations which are superimposed on the high-pressure curve in the combustion chamber (Figure 2). These vibrations are best detected directly in the combustion chamber by means of pressure sensors. As fitting these pressure sensors in the cylinder head for each cylinder is still relatively costly, these vibrations are usually picked up using knock sensors fitted to the exterior of the engine. These piezo-electric acceleration sensors (Figure 7) pick up the characteristic vibrations of knocking combustion and convert them into electrical signals.

There are two types of knock sensor. A wide-band sensor, with a typical frequency band of 5 to 20 kHz, and a resonance sensor, which preferably transmits only one knock-signal resonant frequency. When combined with the flexible signal-evaluation system in the control unit, it is possible to evaluate different or several resonant frequencies from one wide-band knock sensor. This improves knock detection performance, which is why the wide-band knock sensor is increasingly replacing the resonance sensor.

To ensure sufficient knock detection in all cylinders and across all operating ranges, the number and location of the required knock sensors must be carefully determined for each engine type. Four-cylinder in-line engines are usually fitted with one or two knock sensors, while 5- and 6-cylinder engines are fitted with two, and 8- and 12-cylinder engines with four knock sensors.

Signal evaluation
For the duration of a timing range in which knock can occur, a special signal evaluation circuit in the control unit evaluates from the wide-band signal the frequency band(s) with the best knock information and generates a representative variable for each combustion process. This very flexible signal evaluation of the wide-band sensor enables high detection quality. When using a resonance knock sensor that transmits just one resonant frequency for analysis of all cylinders assigned to it across the entire engine map, knock detection is normally no longer possible at higher engine speeds.

Knock detection
The variable produced by the signal-evaluation circuit is classified in a knock-de-

Figure 7: Knock sensor
1 Seismic mass, 2 Sealing compound,
3 Piezoceramic, 4 Connection,
5 Electrical connection.

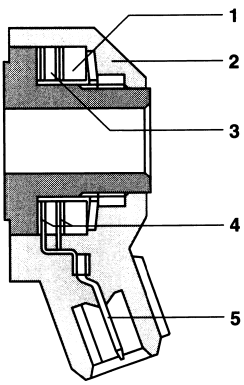
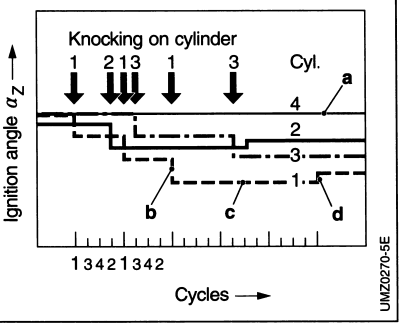


Figure 8: Knock control
Control algorithm for ignition adjustment on a 4-cylinder engine.
$K_{1\ to\ 3}$ Occurrence of knocking at cylinders 1 to 3, at cylinder 4 no knocking.
a Map ignition angle,
b Ignition retard step with knock,
c Ignition advance waiting time,
d Ignition advance step.



tection algorithm as "knock" or "no knock" for each cylinder and for each combustion process. This is done by comparing the variable for the current combustion process with a variable which represents combustion without knock.

Ignition-angle control system with adaptation facility

If combustion knock is detected in a cylinder, the moment of ignition for that cylinder is retarded (Figure 8). When knock stops, the moment of ignition is advanced again in stages up to the precontrol value. The knock-detection and knock-control algorithms are matched in such a way as to eliminate any knock that is audible and damaging to the engine, even though each cylinder is operated at knock limit within the optimum efficiency range.

Real engine operation produces different knock limits, and thus different moments of ignition for individual cylinders. In order to adapt precontrol values for the moment of ignition to a particular knock limit, the ignition retard values are stored for each cylinder dependent on the operating point. They are stored in non-volatile program maps in the permanently powered RAM for load and engine speed. In this way, the engine can be operated at optimum efficiency for each operating point and without audible combustion knocks, even if there are rapid load and engine-speed changes.

This adaptation even enables the use of fuels with lower antiknock properties (e.g. premium instead of Super Plus gasoline).

Phenomenon of pre-ignition

Basic principles

There is currently a clear trend in the development of modern gasoline engines towards downsizing (displacement reduction at same engine performance) and downspeeding (speed reduction by means of longer transmission ratio) in combination with direct injection and supercharging. Supercharging allows for a reduction in displacement without lowering performance level. The engine can therefore be operated in part load at higher loads with higher part-load efficiency, and fuel consumption can be reduced. The charge-air pressure increase to improve efficiency is however limited by the phenomenon of pre-ignition.

A few years ago the terms "extreme knocker" or "superknocker" were used, but they described the symptom and not the cause of the phenomenon of pre-ignition. However, the term "pre-ignition" has become the established term for some time now. In the name there is therefore also a clear distinction from normal combustion knock.

Pre-ignition

Pre-ignitions are uncontrolled auto-ignitions of the air-fuel mixture, which occur before the ignition that is triggered by the spark plug. Compared to normal combustion, this premature initiation of combustion leads to a considerable increase in pressure and temperature, which then leads to severe combustion knock in the subsequent combustion process. In doing so, individual events can pre-damage the engine.

Detection of pre-ignitions

Reliably detecting pre-ignitions is absolutely necessary in order to prevent engine damage. Knocking events as a result of pre-ignitions can be clearly evaluated and detected based on knock sensors by means of position and frequency ranges. This process allows for a high level of detection quality with a clear distinction from normal combustion knock.

Measures to prevent pre-ignitions
If pre-ignitions are detected, countermeasures are necessary in order to prevent further pre-ignitions. However, since pre-ignition is an uncontrolled auto-ignition, there is no direct manipulated variable – e.g. the ignition angle at combustion knock – that can be used to reliably prevent further pre-ignitions. For this reason, the pre-ignition functionality involves initiating multiple countermeasures after detected pre-ignitions in order to quickly reduce the temperature level in the combustion chamber and thus prevent further pre-ignitions. The various measures, such as mixture enrichment or charge lowering, can be activated in the combination that is ideal for the respective engine type.

Pre-ignitions are heavily dependent on fuel and on oil emissions. This functionality is therefore absolutely necessary to protect the engine, especially for turbocharged gasoline engines. This enables vehicle manufacturers to develop efficiency-optimized, supercharged gasoline engines and to market them worldwide despite different fuel and oil grades.

Risk of accidents
All electronic ignition systems are high-voltage systems. To avoid any risks, always switch off the ignition, or disconnect the power supply when working on the ignition system. Such work includes, for example:
- Replacing parts such as spark plugs, ignition coils, or ignition transformers, high-tension ignition cables, etc.
- Connecting engine analyzers, such as stroboscopic lamps, dwell angle/rotational speed testers, ignition oscilloscopes, etc.

When checking the ignition system, remember that dangerously high levels of voltage are present within the system whenever the ignition is on. All tests and inspections should therefore only be carried out by qualified professional personnel.

Ignition systems

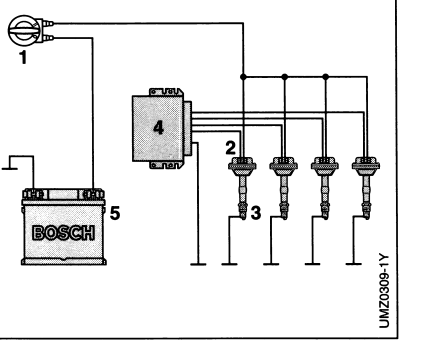
On modern vehicles, the ignition systems are almost always incorporated as subsystems of the engine-management system. Autonomous ignition systems are now only used for special applications (e.g. small engines). In the case of ignition systems, coil ignition (inductive ignition) with a separate ignition circuit per cylinder (static high-voltage distribution with single-spark coils, Figure 9) has come to the fore.

Alongside this, but to a much lesser extent, high-voltage capacitor ignitions (capacitive ignition) or other special designs such as magnetos are used for small engines. The next section will focus on coil ignition alone.

Coil ignition (inductive ignition)

Principle of coil ignition
The ignition circuit of a coil-ignition system (Figure 10) consists of:
- An ignition coil with a primary and a secondary winding
- An ignition output stage to control the current by means of the primary winding, predominantly integrated as an IGBT (Insulated Gate Bipolar Transistor) in the engine control unit or in the ignition coil
- A spark plug connected to the high-voltage connection point of the secondary winding

Figure 9: Ignition system with single spark coils
1 Ignition lock, 2 Ignition coil, 3 Spark plug, 4 Control unit, 5 Battery.



Before the desired moment of ignition, the ignition output stage switches a current from the vehicle electrical system through to the primary winding of the ignition coil. While the primary current circuit is closed (dwell period), a magnetic field builds up in the primary winding.

At the moment of ignition, the current through the primary winding is interrupted again, and the magnetic-field energy is discharged, mainly via the magnetic-coupled secondary winding (induction). In the process, a high voltage is produced in the secondary winding. Flashover occurs if the ignition voltage supply of the ignition system exceeds the ignition-voltage demand of the spark plug. After flashover, the remaining energy is converted at the spark plug while the spark is present.

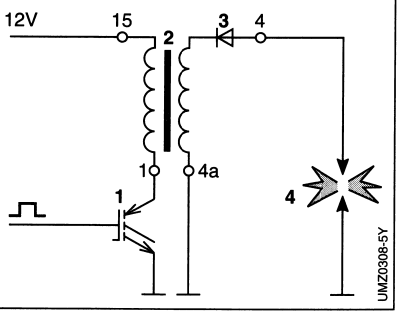
Functions of an ignition system with coil ignition

Moment-of-ignition determination

The current moment of ignition in each case is determined from program maps dependent upon the operating point and output.

Figure 10: Structure of an ignition circuit with single spark coils

1 Ignition output stage,
2 Ignition coil with primary and secondary windings,
3 Activation arc diode (Activation Spark Suppression),
4 Spark plug.
15, 1, 4, 4a Terminal designations,
⎍⎍ Triggering signal.



Dwell-period determination

The required ignition energy is made available at the moment of ignition. The amount of ignition energy is dependent upon the amount of primary current at the moment of ignition (cutoff current) and the inductance of the primary winding. The amount of cutoff current is mainly dependent upon the operating time (dwell period) and upon the battery voltage at the ignition coil. The dwell periods required to achieve the desired cutoff current are contained in characteristic curves or program maps as a function of the steady-state voltage. The change in dwell period with temperature can also be compensated for.

Ignition release

The ignition release ensures that the ignition spark occurs at the right cylinder at the right time and with the required level of ignition energy. On electronic-controlled systems, a trigger wheel with a fixed-angle reference mark (typically 60 – 2 teeth, two teeth missing at the reference mark) on the crankshaft is usually scanned (sensor system with hall or inductive sensor). From this, the control unit can calculate the crankshaft angle and the momentary rotational speed. The ignition coil can be switched on and off at any required crankshaft angle. An additional phase signal from the camshaft is required for the unambiguous identification of the cylinder.

For each combustion, the control unit uses the desired moment of ignition, the required dwell period and the current engine speed to calculate the switch-on time and switches on the output stage. The moment of ignition, or the switchoff point for the output stage, can be triggered either when the dwell period expires or when the desired angle is reached.

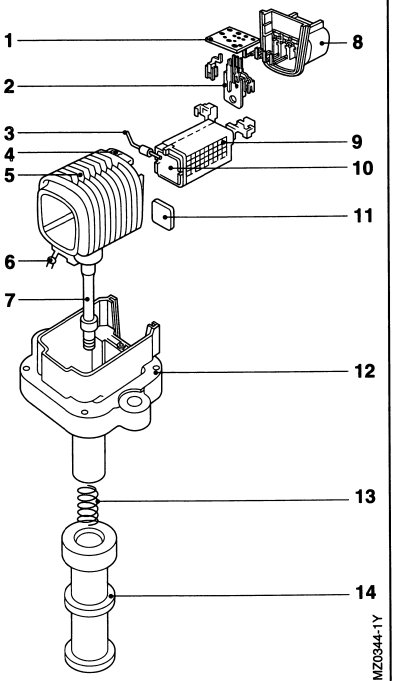
Ignition coil

Function

The ignition coil is principally an energy-charged high-voltage source similar in structure to a transformer. Energy is supplied by the vehicle electrical system during the dwell period or charging time. At the moment of ignition, which at the same time is the end of the charging time, the energy is then transferred with the required high voltage and sparking energy to the spark plug (inductive ignition system).

Figure 1: Design of the compact ignition coil

1 Printed-circuit board (optional),
2 Ignition output stage (optional),
3 Activation arc diode (optional),
4 Secondary-winding housing,
5 Secondary winding, 6 Contact plate,
7 High-voltage pin
(connecting element to contact spring),
8 Connector, 9 Primary winding,
10 I-core, 11 Permanent magnet, 12 O-core,
13 Contact spring (spark-plug contacting),
14 Silicone jacket (high-voltage insulation).



Design

The ignition coil (Figure 1) comprises two windings that are magnetically linked by an iron core (I-core and O-core). This iron core may contain a permanent magnet for energy optimization. Compared with the secondary winding, the primary winding has significantly fewer coils. The turns ratio $\ddot{u}$ is in the range of $\ddot{u}$ = 80 to 150.

The windings must have good electrical insulation to prevent electrical discharge and flashovers either to the inside or to the outside. For this purpose, the windings are usually cast in epoxy resin in the ignition coil housing.

As a rule, the iron core consists of stacked, ferromagnetic steel-plate fins in order to minimize above all eddy-current losses.

Alternatively, the ignition output stage can be incorporated into the ignition coil instead of the engine control unit. Interference-suppression elements may also be incorporated into the ignition coil together with the activation arc diode (activation spark suppression). An interference-suppression resistor is commonly used on the high-voltage output to the spark plug.

Operation

The ignition output stage switches the primary current in the ignition coil. The current rises with a delay, in accordance with the inductance. Energy is stored in the magnetic field created in the process. The operating time (charging time) is calculated in such a way that a specific cutoff current, and thus a specific level of stored energy, will be reached by the end of the operating time.

The current is cut off by the ignition output stage, which causes a swift change in the magnetic flux in the ignition coil's iron core. This change in flux results in a voltage induction in the secondary winding. The design of the secondary winding, the geometric arrangement of the secondary winding to the iron core and the primary winding and the materials used in the insulation system and the iron core give rise to inductive and capacitive properties that lead to a voltage supply in excess of 30,000 V on the secondary side.

When the ignition-coil voltage supply equals the ignition-voltage demand at the spark-plug electrodes, the voltage drops to

a spark voltage of about 1,000 V. A spark current then flows and decreases as spark duration increases, until the spark is finally extinguished.

As the current changes over time when the primary current is cut off, an induction voltage occurs at the ignition coil output, similar to the start of charging time. However, this voltage is much smaller than the voltage at the moment of ignition and has an inverse polarity to that voltage. To prevent this "transient voltage" causing unwanted ignition, it is usually suppressed by means of a high-voltage diode in the secondary circuit (activation arc diode, activation spark suppression).

The design of an ignition coil can determine its electrical characteristics. In this respect, the requirements in terms of installation space (e.g. geometry of cylinder head and cylinder-head cover, position of fuel injector and intake manifold) and the two specified interfaces, the ECU with output stage (e.g. cutoff current) and the spark plug (e.g. ignition voltage, spark data), are decisive.

Design variations

There are several types of ignition coil that are distinguishable by their various features.

Individual coils and modules

In addition to the individual coils, which normally sit directly on the spark plug, several ignition coils can be grouped together in a single housing as a module. They are then mounted directly on the spark plugs, or they may be located a short distance away, in which case the high voltage must be supplied via appropriate cables.

Single-spark and dual-spark ignition coils

In addition to ignition coils with just one high-voltage output (single-spark ignition coil), there are also coils that use both ends of the secondary winding as an output (dual-spark ignition coil). The electric circuit must always be closed by both spark gaps to prevent discharge on the secondary side. One feasible application for this design is dual ignition, i.e. two spark plugs per cylinder supplied by one ignition coil. Another application involves splitting up the two high-voltage outputs between two spark plugs on different cylinders. In this case, one of the two spark plugs will always be in the ignition stroke. As a result, the voltage and energy requirements of the "passive spark" (backup spark) are reduced substantially. Above all, this variant offers cost advantages; however, it must be matched to the overall system to prevent damage from unwanted ignition events caused by backup sparks.

Compact coil and pencil coil

Ignition coils are also distinguishable by their basic design. For example, there is the conventional compact coil, that features an equal-sided coil body and an O-I core or C-I core magnetic circuit. The coil body sits in the engine above the spark-plug well.

Another type is the pencil coil whose coil body projects into the spark-plug well. Here, too, the windings are located on an I or pencil core, with a plate (yoke plate) arranged concentrically around the windings serving as a magnetic yoke.

Ignition coils with ignition output stage

Ignition coils are available with or without the ignition output stage. The reason for integrating the output stage is to relieve the strain on the engine ECU. Additionally, further electronic circuitry is used in the ignition coil in order to meet the existing requirements (Figure 2).

Requirements

The principal requirements of modern ignition systems are derived indirectly from necessary emission and fuel reductions. Requirements of ignition coils are derived from relevant engine solutions, such as high-pressure supercharging and lean-burn and stratified-charge operation (spray-guided direct injection) in combination with increased exhaust-gas recirculation rates (EGR).

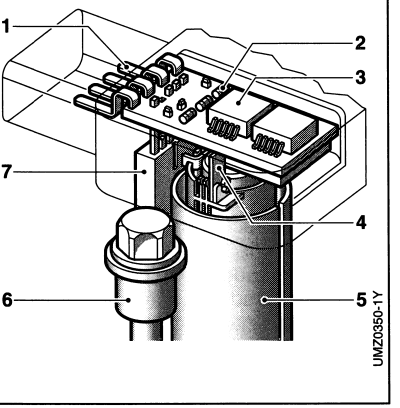
It is necessary to represent an increased ignition-voltage and energy demand with increased temperature requirements. This is effected among others by means of

- High-energy coils with high voltage supply (in excess of 40,000 V)
- Multispark ignition (MSI), i.e. spark-sequence control with integrated electronics (ASIC), controlled by way of primary- and secondary-current evaluation
- Diagnostic functions (e.g. charge-time monitoring, ionic-current measurement for combustion diagnosis)
- Protective functions (e.g. overtemperature cutoff, cutoff-current regulation)

Furthermore, motor vehicles are subject to increased demands with regard to electromagnetic compatibility (EMC). Precisely because of the increased ignition-voltage demand, the increased ignition frequencies necessitated by multispark ignition and the increased cutoff currents, it is now necessary to reduce the interference emitted by the ignition system so as not to disrupt the functioning of other automotive components (ECUs, microcontrollers, sensors, actuators).

Figure 2: Installation of ignition output stage in a pencil ignition-coil housing

1 Connector,
2 SMD components (Surface Mounted Device),
3 Electronic circuitry for ignition functions,
4 Primary-winding contacts,
5 Yoke plate of pencil ignition-coil transformer,
6 Attachment lug,
7 Ignition output stage.



The integration of increasingly more electronic circuitry in ignition coils gives rise to more stringent demands with regard to the interference immunity of these components. In the interests of avoiding impaired functioning and malfunctions, the electronic circuitry must be immune to the emissions of the ignition system itself and to the interference emitted by other vehicle components.

Spark plug

Function

The spark plug transfers the ignition-coil energy into the combustion chamber in a spark-ignition engine. The high voltage applied creates an electric spark between the spark-plug electrodes, which ignites the compressed A/F mixture and thereby initiates combustions.

In combined operation with the other engine components, e.g. the ignition and the mixture-formation systems, the spark plug determines the function of the spark-ignition engine to a decisive extent. It must provide for reliable cold starting, guarantee misfire-free operation over its entire service life (at least 30,000 km), and not exceed the maximum permissible temperature even during extended operation at critical operating points. Furthermore, the spark plug must always reliably insulate the high voltage against the cylinder head and seal the combustion chamber towards the outside.

Requirements

The spark plug must satisfy a variety of challenging performance demands. It is exposed to the varying periodic processes within the combustion chamber, as well as external climatic conditions.

When spark plugs are used with electronic ignition systems, ignition voltages of up to 40 kV can occur; these high voltages must not cause ceramic or insulator-head flashovers. The deposited residues from the combustion process such as soot, carbon and ash from fuel and oil additives are electrically conductive under certain thermal conditions. Nevertheless, no flashovers via the insulator are permitted to occur. The electrical resistance of the insulator must be sufficiently high up to 1000 °C and is only permitted to drop slightly over the spark plug's service life.

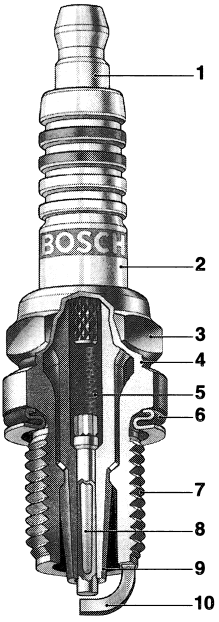
Mechanically, the spark plug is subjected to the pressures (up to 150 bar) occurring periodically in the combustion chamber; however, gas tightness may not be impaired. In addition, the spark-plug electrode materials exhibit extreme resistance to thermal loads and continuous vibratory stress. The spark-plug shell must be able to absorb the forces occurring during assembly without any lasting deformation.

At the same time, the section of the spark plug that protrudes into the combustion chamber is exposed to high-temperature chemical processes, making resistance to aggressive combustion deposits essential.

Above all, the spark-plug insulator and the electrodes must be capable of withstanding high thermal loads as they are exposed in rapid succession to high temperatures from the hot combustion gases and low temperatures of the cold air/fuel mixture (thermal shock). The electrodes and the insulator at the cylinder head must have good heat dissipation properties – essential for reliable spark-plug perfor-

Figure 1: Spark-plug design

1 Terminal stud
(here with SAE terminal),
2 Insulator head,
3 Nickel-plated steel shell,
4 Heat-shrinkage zone,
5 Electrically conductive glass seal,
6 Captive gasket (seal seat),
7 Thread,
8 Compound center electrode (Ni/Cu),
9 Insulator nose,
10 Ground electrode
(here Ni/Cu compound electrode).



mance. The connection end of the spark plug should heat up as little as possible.

As well as these requirements, the spark plug must also be adapted to the geometric specifications of the engine design (e.g. spark-plug position in the cylinder head).

In view of these requirements – resulting from the great variety of engines – a wide range of spark plugs is required.

Design

The spark plug consists of at least two electrodes (center and ground electrodes), between which the ignition spark is generated. The high voltage is transferred to the center electrode via the terminal stud by an electrically conductive glass seal inside the insulator. The insulator prevents a short-circuit to the spark-plug shell, in which the plug (insulator with operationally reliable mounted center electrode and terminal stud) is mounted (Figure 1).

Terminal stud and high-voltage terminal

The steel terminal stud has at the end projecting out of the insulator a thread (M4) into which the spark-plug connector of the ignition cable engages. For the standardized connector either a terminal nut (SAE) is screwed onto the terminal-stud thread or the stud is already provided during manufacture with a solid standardized terminal (e.g. cup). Spark plugs with metal shields are available for watertight systems and for maximum interference suppression.

Conductive glass seal

This electrically conductive glass seal acts as a mechanical support for the components, while providing a gas seal against the high combustion pressure. It can also incorporate resistor elements for interference suppression and burn-off.

Insulator

The insulator consists of a special ceramic. Its function is to insulate the center electrode and the terminal stud electrically against the spark-plug shell. The requirements of good thermal conductivity with simultaneously high insulation resistance are in stark contrast to the properties of most insulating materials. The material used by Bosch consists of aluminum oxide (Al_2O_3), to which other substances are added in small quantities.

At the connection end, the insulator has a lead-free glaze to repel moisture and dirt. This avoid insulator flashovers to a great extent.

Spark-plug shell

The shell is made of steel. The lower part of the shell has a thread so that the spark plug can be secured in the cylinder head and exchanged after a specified change interval. Depending on the spark-plug concept, up to four ground electrodes are welded onto the end face of the shell.

To protect the shell against corrosion, its surface is nickel-plated so as to prevent the thread from seizing in the aluminum cylinder heads. The upper part of the shell features a hexagon or bi-hexagon head for application of a wrench. The bi-hexagon head requires less space in the cylinder head while the insulator-head geometry remains unchanged, allowing the engine design engineer greater freedom in designing the cooling ducts.

The upper part of the shell is flanged after the plug is fitted and locates the latter in its position. The subsequent shrinking process by inductive heating under high pressure creates the gastight connection between the insulator and the shell, and guarantees good heat conduction.

Seal seat
Depending on the engine design, the spark plug has a flat or conical seal seat, which provides the seal against the cylinder head. In the case of the flat seal seat, a "captive" sealing ring is used as the sealing element. It has a special shape and provides a permanently elastic seal when the spark plug is fitted. In the case of the conical seal seat, a conical shell surface provides a direct seal on a corresponding surface of the cylinder head without the use of a sealing ring.

Electrodes
The electrodes subjected to high thermal load are primarily made of a nickel-based multi-alloy. Through the addition by alloying of manganese and silicon the chemical resistance of nickel is improved, above all against the highly aggressive sulfur dioxide (SO_2, sulfur is an ingredient of lube oil and fuel). Additives made of aluminum and yttrium additionally increase the resistance to scaling and oxidation.

Compound electrodes with a jacket material made of a nickel alloy and a copper core are used to improve heat dissipation and consequently wear performance.

Center electrode
The center electrode is anchored with its head in the conductive glass seal. For longlife spark plugs it is sensible to use corrosion- and oxidation-resistant materials like platinum (Pt) and platinum alloys, which also exhibit high burn-off resistance. The center electrode then accommodates a precious-metal pin, which is permanently joined to the base electrode by laser welding.

Ground electrodes
The ground electrodes are welded to the shell and have predominantly a rectangular cross-section. Depending on the arrangement, a distinction is made between front and side electrodes (Figure 3) and special applications (e.g. spark plug without distinct ground electrode for racing engines). As well as thermal conductivity, the length, the profile cross-section and the number of ground electrode determine their temperature and hence their wear performance.

Heat range
Operating temperature of the spark plug
When the engine is operating, the spark plug is heated by combustion heat. Some of the heat absorbed by the spark plug is diverted to the fresh A/F mixture. Most of the heat is transmitted to the spark-plug shell via the center electrode and the insulator, and is diverted to the cylinder head. The operating temperature of the spark plug represents a balance between heat absorption from the engine and heat dissipation to the cylinder head.

The aim is for the insulator nose to reach a self-cleaning temperature of approx. 500 °C even at low engine performance. If the temperature drops below this level, there is the danger that soot and oil residue from incomplete combustion will settle on the cold areas of the spark plugs (particularly when the engine is not at normal operating temperature, at low outside temperatures and during repeated starts) (Figure 2, curve 3). This can create a conductive connection (shunt) between the center electrode and the spark-plug shell. This will cause ignition energy to leak away in the form of short-circuit current (risk of misfiring).

At higher temperatures, the residue containing carbons burn on the insulator nose; the spark plug thus "cleans" itself (Figure 2, curve 2).

An upper temperature limit of approx. 900 °C should be observed since, in this range, spark-plug electrode wear increases drastically (due to oxidation and hot-gas corrosion). If this limit is exceeded by a significant extent, it increases the risk of auto-ignition (ignition of the A/F mixture on hot surfaces) (Figure 2, curve 1). Auto-ignition subjects the engine to extreme loads, and may result in engine destruction within a short period of time. The spark plug must therefore be adapted to the engine type in terms of its heat-absorbing property.

Heat-range code number
The identifying feature of a spark plug's thermal loading capacity is its heat range, which is defined by a heat-range code number and determined in comparison measurements with a reference standard source (calibration spark plug).

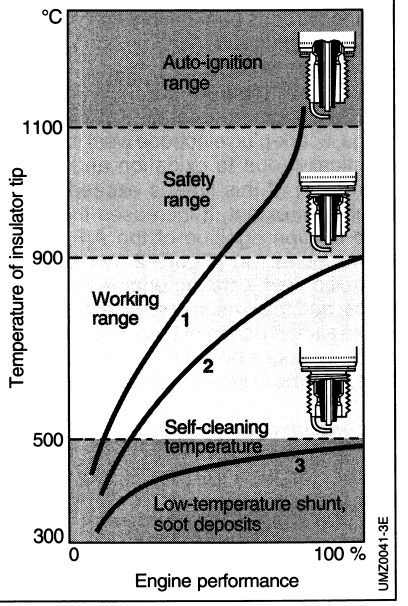
A low code number (e.g. 2 to 5) denotes a "cold spark plug" with low heat absorption through a short insulator nose. High heat-range code numbers (e.g. 7 to 10) denote "hot spark plugs" with high heat absorption through long insulator noses.

The heat range matching the respective engine is determined in the course of application measurements during development.

Ionic-current measurement procedure
The Bosch ionic-current measurement procedure employs the combustion process as a factor for determining the heat-range requirement. The ionizing effect of flames is used to assess how combustion develops over time; this is done by measuring conductivity in the spark gap.

Figure 2: Spark-plug temperature response
1 Spark plug with excessively high heat-range code number (hot plug),
2 Spark plug with suitable heat-range code number,
3 Spark plug with excessively low heat-range code number (cold plug).

The working range temperature should be from 500 °C to 900 °C for different engine performance ratings.



Characteristic changes in the combustion process due to increased thermal loading of the spark plugs can be detected using ionic current and used in the assessment of the auto-ignition process.

Thermal ignition
Ignitions of the air/fuel mixture which occur independently of the ignition spark and for the most part on a hot surface (e.g. on the excessively hot insulator-nose surface of a spark plug with excessively high heat range) are called auto-ignitions. These can be subdivided into two categories on the basis of their position in time relative to the moment of ignition.

Post-ignition
Post-ignitions occur after the electrical moment of ignition, but are not critical for practical engine operation since electrical ignition occurs at an increasingly earlier stage. To find out whether thermal ignitions are initiated by the spark plug, individual ignitions are cyclically suppressed during the ionic-current measurement. When a post-ignition occurs, the ionic current increases significantly only after the moment of ignition. But since a combustion is initiated, a pressure increase and hence a torque output are also to be observed.

Pre-ignitions
Pre-ignitions occur before the electrical moment of ignition and can, on account of their uncontrolled progression (see Irregular operating states), cause serious engine damage. Excessively early combustion initiation shifts not only the position of the pressure maximum to top dead center (TDC), but also the maximum combustion-chamber pressure to higher values. This increases thermal and mechanical load on the components in the combustion chamber.

Spark-plug choice
The aim of an adaptation is to choose a spark plug which can be operated without thermal ignitions before the moment of ignition (pre-ignition) and has a sufficient heat-range reserve. In this way, the variations in engine and spark-plug production are covered and it is also taken into consideration that engines can change in terms of

their thermal properties over the course of their operating life. Thus, for example, oil-ash deposits in the combustion chamber can increase the compression ratio, which in turn causes a higher temperature load on the spark plug.

The correct heat range for the engine is determined when no failures with sooted spark plugs occur in the concluding cold-start tests with this heat-range recommendation.

The use of materials with higher thermal conductivity for the center electrodes (silver or nickel alloys with copper core) provides – with the same heat-range code number – for a significant increase in the insulator-nose length, the upshot of which is that the spark plug is less prone to heavy sooting. This reduces the likelihood of combustion misses and misfiring – with their attendant massive increases in hydrocarbon emissions – and provides benefits in exhaust emissions and fuel consumption in part-throttle operation at low load factors.

Close cooperation between engine and spark-plug manufacturers is customary when it comes to choosing suitable spark plugs.

Electrode gap and ignition voltage

The electrode gap is the shortest distance between the center electrode and the ground electrode (Figure 3). It determines, among other things, the ignition-spark length and should on the one hand be as large as possible to enable the ignition spark to activate a large area of the air/fuel mixture. This gives rise through stable flame-core formation to a reliable ignition of the air/fuel mixture. On the other hand, a lower ignition voltage is needed to produce a spark with a smaller electrode gap. If the electrode gap is too small, however, only a small flame core will be created around the electrode. Since energy will be drawn from the flame core via the contact areas with the electrodes (quenching), the flame core will only be able to propagate very slowly. In extreme cases, so much energy may be drawn off that misfires will occur.

During flashover and operation at higher temperature the electrode material is subjected to such high stress that the electrodes wear and consequently the electrode gap gets bigger. As the electrode gap increases the ignition conditions are in fact improved, but the ignition-voltage demand increases. As the ignition coil's voltage supply is a fixed amount, the ignition-voltage reserve is reduced and the risk of misfiring increases.

The required ignition voltage is influenced not only by the electrode gap, the electrode shape, the temperature, and the electrode material, but also by parameters specific to the combustion chamber, such as mixture composition (λ value), flow velocity, turbulence, and density of the ignitable gas.

In modern engine concepts with high mixture density, which frequently feature high charge turbulence, electrode gaps must be carefully defined in order to guarantee reliable ignition, and thus misfire-free operation throughout the required service life.

Spark position

The position of the spark gap relative to the combustion chamber wall defines the spark position. On modern engines (and particularly with gasoline direct-injection engines), the spark position has a considerable influence on combustion. A perceptible improvement in ignition response is observed when the spark position projects deeper into the combustion chamber. Combustion can be characterized by smooth-running engine performance, which can be derived directly from engine speed fluctuations.

However, because the ground electrodes are longer, higher temperatures are achieved. This, in turn, has an effect on electrode wear and electrode durability. It is possible to achieve the required service life by implementing design measures (e.g. extending the spark-plug shell beyond the combustion chamber wall), or by using compound electrodes, or high-temperature-resistant materials.

Spark-plug concepts

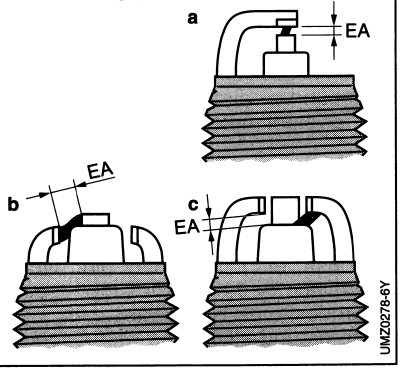
One or more ground electrodes can be advantageous, depending on the demands placed on the spark plug (wear, ignition response, etc.). The spark-plug type is determined by the relative location of the electrodes to each other and the position of the ground electrodes with respect to the insulator.

Air-gap spark concept
In the air-gap spark concept (Figure 3a) the ignition spark crosses between the center electrode and the ground electrode by the direct route and ignites the air/fuel mixture.

Surface-gap spark concept
The position of the ground electrode relative to the ceramic produces the surface-gap spark concept, in which the spark travels from the center electrode across the surface of the insulator nose before jumping across a gas-filled gap to the ground electrode (Figure 3b). The advantages here are the lower voltage demand with the same electrode gap, improved ignition properties,and the insulator-cleaning effect of the surface-gap spark for better repeat cold-starting performance.

Surface-air-gap spark concept
When both forms of discharged are made possible by the specific geometric arrangement of the ground electrodes, this is called the surface-air-gap spark concept (Figure 3c). Depending on operating conditions, the spark behaves either as an air-gap spark or a surface-gap spark with different ignition-voltage demand values.

Figure 3: Spark-plug concepts
a) Air-gap spark concept with front electrode,
b) Surface-gap spark concept with side electrodes,
c) Surface-air-gap spark concept with side electrodes.
EA Electrode gap.



In view of the trend towards increasing combustion-chamber pressures, air-gap spark concepts are to be preferred since these do not cause spark embedding in the insulator.

Simulation-based development of spark plugs
The Finite-Element Method (see FEM examples) is used for spark plugs to calculate temperature and electric fields and to solve structure-mechanical problems. Changes to the geometry and materials of spark plugs or even different physical boundary conditions and their effects can thus be determined in advance without complicated tests. This is the basis for a targeted production of test specimens, with which the calculation results are verified by way of example.

Spark-plug operating performance
Electrode wear
As spark plugs operate within an aggressive atmosphere, sometimes at extremely high temperatures, electrodes are subject to wear. This material erosion causes the electrode gap to increase noticeably the longer the spark plug is in service, therefore causing the ignition-voltage demand to increase. When this requirement can no longer be met by the supply from the ignition coil, misfiring will occur.

There are essentially two mechanisms that are responsible for electrode wear: spark erosion and corrosion in the combustion chamber. The flashover of electric sparks results in an increase in the temperature of the electrodes up to their melting temperature. The fused microscopically small surface areas react with the oxygen or the other constituents of the combustion gases. The consequence is material erosion.

Materials with a high thermal resistance (special-steel alloys of platinum or iridium) are used to minimize electrode wear. But material wear can also be reduced for the same period of use by the appropriate selection of electrode geometry (e.g. smaller diameters, thin pins) and spark-plug concept (surface-gap spark plugs). In a surface-air-gap spark plug with four ground electrodes, there are eight possible spark gaps. In this way, the wear is evenly distributed across all four electrodes.

The electrical resistance effected in the conductive glass seal also reduces erosion and wear.

Changes in operation
Dirt and changes in the engine caused by aging (e.g. higher oil consumption) can also affect operation of the spark plug. Deposits on the spark plug can result in shunts, and thus in misfiring. This, in turn, may cause a considerable rise in pollutant emissions, and even damage the catalytic converter.

For these reasons, spark plugs have a defined service life after which the spark plugs must be replaced.

Irregular operating states
Irregular operating states may arise and the engine and the spark plugs may sustain damage due to incorrectly set ignition systems, the use of spark plugs whose heat range is unsuitable for the engine, or the use of unsuitable fuels.

Auto-ignition
Auto-ignition (pre-ignition) is an unwanted ignition process. This arises in response to overheated components (e.g. on the tip of the spark-plugs insulator nose, on an exhaust valve, on protruding cylinder-head gaskets) with the result that the air/fuel mixture ignites there under uncontrolled conditions. Severe damage can be caused to the engine and spark plug as a result of auto-ignition.

Combustion knock
Knocking is characteristic of an uncontrolled combustion process with very sharp rises in pressure (see Knocking). The combustion process is considerably faster than normal combustion. The components (cylinder head, valves, piston, and spark plugs) are subject to high temperature load which can result in damage. Knocking can be prevented by retarded ignition angle (see Knock control).

Versions and applications
The different spark-plug types are identified by a type designation. This contains all the essential spark-plug features (Figure 4, [1]). The electrode gap is additionally indicated on the packaging. The spark plug which is suitable for a given engine is specified or recommended by the engine manufacturer and by Bosch.

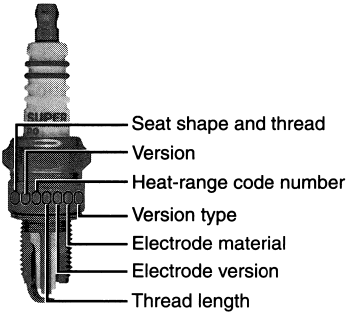
Standard spark plug for automobiles
The spark plug pictured in Figure 1 can be referred to as the standard plug for older, simply built, naturally aspirated engines.

The demands imposed on the spark plug are constantly increasing as engines are further developed up to higher specific power outputs due to, for example, improved fuel efficiency and emission-control legislation. The most sophisticated engines are direct-injection engines with multi-stage supercharging and valve-timing adjustment.

Spark plugs for direct-injection engines
Direct-injection engines impose particularly high demands on spark plugs. They must therefore be specially adapted to the needs of the respective engine (e.g. power output, mean pressure). This results in different requirements of the spark-plug concept for combustion processes with stratified or homogeneous operation.

Through the combination of direct injection with turbocharging, the emphasis in spark-plug development is on expanding

Figure 4: Designation code
The type designation describes the following features.



ignition-voltage demand, thermal load capability of the electrodes, mechanical load capability, and service life (wear reduction). For this reason, spark plugs with precious metal electrodes and M12 thread are normally used, and they are also specifically aligned in the combustion chamber when used for spray-guided combustion processes with stratified operation.

Spark plugs for gas and flex-fuel applications
The spark-plug design for natural-gas and flex-fuel use is comparable with that described previously. For each application the appropriate spark-plug version is determined while taking into account the operating conditions within the framework of the application measurements.

As a rule, alternative fuels have higher antiknock properties such that the moment of ignition can be advanced for an optimum combustion. For the most part a lower heat range is recommended for the spark plug in view of the resulting higher thermal load on the spark plug.

Special spark plugs
Special spark plugs are used for particular requirements. These differ in their design, which is dictated by the usage conditions and the installation conditions in the engine.

Spark plugs for motor sport
Engines for sports vehicles are exposed to extreme thermal loads on account of the high proportion of wide-open throttle applications involved. Spark plugs for these operating conditions mostly have precious-metal electrodes (silver, platinum) and a short insulator nose with low heat absorption.

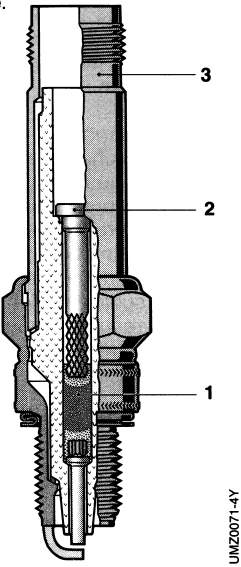
Fully-shielded spark plugs
In the event of very high interference-suppression demands, it may be necessary to shield the spark plugs. On fully-shielded spark plugs the insulator is wrapped in a metal shielding sleeve. The terminal is located inside the insulator (Figure 5). Fully-shielded spark plugs are watertight.

Spark-plug practice
Spark-plug installation
When installed correctly and the right type is chosen, the spark plug is a reliable component of the ignition system. The electrode gap can only be re-adjusted in spark plugs with front electrodes. On surface-gap and surface-air-gap spark plugs the ground electrodes must not be re-adjusted as this will alter the spark-plug concept.

Errors and their consequences
For a particular engine type only the spark plugs approved by the engine manufacturer or recommended by Bosch may be used. Using unsuitable spark-plug types can cause serious engine damage.

Incorrect heat-range code number
The heat-range code number must without fail match the engine manufacturer's spark-plug specification or the recommendation by Bosch. Auto-ignitions may result if spark plugs with a heat-range code number other than that prescribed for the engine are used.

Figure 5: Fully-shielded spark plug
1 Special conductive glass seal (interference-suppression resistor),
2 Ignition-cable terminal,
3 Shielding sleeve.



Incorrect thread length
The thread length of the spark plug must match the thread length in the cylinder head. If the thread is too long, the spark plug will project too far into the combustion chamber. Damage to the piston is a potential consequence. In addition, coked spark-plug threads may make it impossible to unscrew the spark plug or cause the spark plug to overheat.

If the thread is too short, the spark plug will not project far enough into the combustion chamber. This may result in poorer mixture ignition. Furthermore, the spark plug will not reach its self-cleaning temperature and the lower threads in the cylinder head will sustain coking.

Tampering with the seal seat
On spark plugs with conical seal seats neither a plain washer nor a sealing ring may be used. On spark plugs with flat seal seats only the "captive" sealing ring located on the spark plug may be used. It must not be removed or replaced by a plain washer. Without the sealing ring the spark plug projects too far into the combustion chamber, the transfer of heat from the spark-plug shell to the cylinder head is compromised, and the spark-plug seat provides a poor seal. If an additional sealing ring is used, the spark plug does not project far enough into the threaded hole, and the transfer of heat from the spark-plug shell to the cylinder head is likewise compromised.

Assessment of spark-plug faces
The spark-plug face refers to the combustion end of the spark plug with the electrodes and the insulator nose. Their appearance in terms of color and deposits provides indications of the spark plug's operating performance and the engine's mixture composition and combustion process.

References
[1] www.bosch-zuendkerze.de

Catalytic exhaust-gas treatment

Catalytic converter

Emission-control legislation defines limits governing pollutant emissions from motor vehicles. Engine-design measures alone are not sufficient to comply with these limits. In addition to reducing untreated engine emissions, the emphasis in spark-ignition engines is on catalytic aftertreatment of the exhaust gas in order to convert the pollutants. Catalytic converters convert the pollutants produced during combustion into harmless components.

Three-way catalytic converter

Function

The state-of-the-art technology for engines operating with a stoichiometric A/F mixture is the three-way catalytic converter (TWC). Its task is to convert pollutant components HC (hydrocarbons), CO (carbon monoxide), and NO_x (nitrogen oxides), which arise during the combustion process, into harmless components. The end products are H_2O (water vapor), CO_2 (carbon dioxide), and N_2 (nitrogen).

Design and operating principle

The catalytic converter comprises a steel-plate container as housing, the substrate and the substrate coating (washcoat) made of aluminum oxide (Al_2O_3) on which the precious metal is finely spread. The substrate is usually a ceramic monolith, though metallic monoliths are used for special applications (see Ceramic monolith and Metal monolith). On the monolith is a substrate coating, which enlarges the effective area of the catalytic converter by a factor of up to around 10,000. The catalytic layer on top of this contains the precious metals palladium or in the past platinum and rhodium. Platinum and palladium accelerate the oxidation of HC and CO, while rhodium is responsible for reducing NO_x. The amount of precious metal contained in a catalytic converter is approx. 1 to 10 g depending on the engine displacement and the exhaust emissions standard to be met.

For example, oxidation of CO and HC takes place according to the following equations:

$$2\,CO + O_2 \rightarrow 2\,CO_2,$$
$$2\,C_2H_6 + 7\,O_2 \rightarrow 4\,CO_2 + 6\,H_2O.$$

Reduction of nitrogen oxides takes place, for example, according to the following equations:

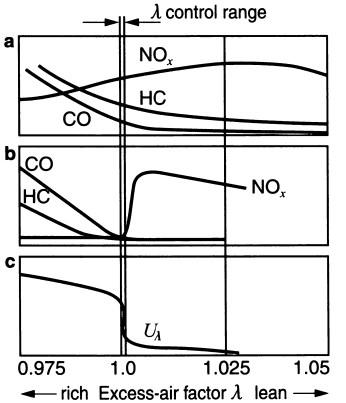
$$2\,NO + 2\,CO \rightarrow N_2 + 2\,CO_2,$$
$$2\,NO_2 + 2\,CO \rightarrow N_2 + 2\,CO_2 + O_2.$$

The oxygen required for the oxidation process is either present in the exhaust gas as exhaust-gas oxygen (resulting from incomplete combustion), or it is taken from the NO_x, which is reduced at the same time.

The concentrations of pollutants in the untreated exhaust gas (upstream of the catalytic converter) depend on the excess-air factor λ (Figure 1a). In order for the conversion of the three-way catalytic

Figure 1: Catalytic-converter efficiency as a function of the excess-air factor λ

a) Exhaust-gas emission upstream of three-way catalyst,
b) Exhaust-gas emission downstream of three-way catalyst,
c) Electrical signal of two-step λ sensor.
U_λ Sensor voltage.



converter to be as high as possible for all three pollutant components, a mixture composition in the stoichiometric ratio $\lambda = 1$ is required (Figure 1b). At $\lambda = 1$, there is a state of equilibrium between oxidation and reduction reactions which facilitates a full oxidation of HC and CO and, at the same time, reduces the NO_x. In this way, HC and CO act as reducing agents for NO_x. The "window" (λ control range) in which the mean time value of λ must lie is very small. Mixture formation must therefore be corrected in a λ control loop with the aid of the signal of a λ oxygen sensor (Figure 1c) (see λ control).

Oxygen storage
The λ accuracy in the dynamic range is typically 5 %, i.e. fluctuations of $\lambda = 1$ in this scale are inevitable. The catalytic converter can compensate for minor mixture fluctuations itself. It has the ability to store excess oxygen in the lean phase in the catalytic converter in order to release it again in the subsequent rich phase. Its substrate coating contains ceroxide, which can make oxygen available via the following balance reaction:

$$2\ Ce_2O_3 + O_2 \leftrightarrow 4\ CeO_2.$$

The task of the engine-management system is, therefore, clear: The mean time value of the resulting λ upstream of the catalytic converter must be very precise at one (a deviation of only a few thousandths is permitted). The mean value deviations, converted to oxygen input and output, must not overtax the available oxygen retained in the catalytic converter. Typical oxygen-storage values are in the range of 100 mg to 1 g; these values decrease as the catalytic converter ages. All conventional methods of diagnosing catalytic converters are based on the direct or indirect determination of this oxygen storage capacity (osc).

NO_x storage catalytic converter
Function, design and operating principle
During lean-burn operation, it is impossible for the three-way catalytic converter to convert the nitrogen oxides (NO_x) which have been generated during combustion. CO and HCs are oxidized by the high residual-oxygen content in the exhaust gas and are therefore no longer available as reducing agents for the nitrogen oxides.

The catalytic layer of the NO_x storage catalytic converter also contains substances which can store NO_x (e.g. barium oxide). All conventional NO_x storage coatings also have the properties of a three-way catalytic converter, with the result that the NO_x storage catalytic converter operates like a three-way catalytic converter at $\lambda = 1$.

In stratified lean-operation mode (see Gasoline direct injection), NO_x is not converted continuously, but instead in three stages.

NO_x storage
During storage, NO_x is initially oxidized to NO_2, which then reacts with the special oxides of the catalytic-converter surface and oxygen (O_2) to form nitrates (e.g. barium nitrate):

$$2\ NO + O_2 \rightarrow 2\ NO_2,$$
$$2\ BaO + 4\ NO_2 + O_2 \rightarrow 2\ Ba(NO_3)_2.$$

Regeneration
As the quantity of stored NO_x (load) increases, the ability to continue binding NO_x decreases. At a predefined laden state, the NO_x accumulator must be regenerated, i.e. the nitrogen oxides stored in it must be released again and converted. For this purpose, the engine switches briefly to rich homogeneous operation ($\lambda < 0.8$) to reduce NO to N_2, without emitting CO and HC in the process. In a second stage, the rhodium coating reduces the nitrogen oxides with CO:

$$Ba(NO_3)_2 + 3\ CO \rightarrow 3\ CO_2 + BaO + 2NO.$$
$$2\ NO + 2\ CO \rightarrow N_2 + 2\ CO_2.$$

The end of the storage and release phase is either calculated using a model-based method or measured with a λ oxygen sensor after the catalytic converter.

Desulfating
The sulfur contained in the fuel also reacts with the accumulator material in the catalytic layer. The result is that, over time, the amount of accumulator material available for NO_x accumulation diminishes. This creates sulfates (e.g. barium sulfate) which are very temperature-resistant and are not reduced during NO_x regeneration. For desulfating, special measures must be used to heat the catalytic converter to 600 to 650 °C before exposing it alternately to rich (λ = 0.95) and lean exhaust gas (λ = 1.05) for a few minutes. This process reduces the sulfates.

The various methods to heat up the NO_x storage catalytic converter located in the underfloor regions must be careful not to overheat the close-coupled catalytic converter.

Catalytic-converter operating temperature
Catalytic converters do not start significant conversion until they have reached a specific operating temperature (light-off temperature). On a three-way catalytic converter, this is approximately 300 °C. Ideal conditions for a high conversion rate are achieved at 400 to 800 °C.

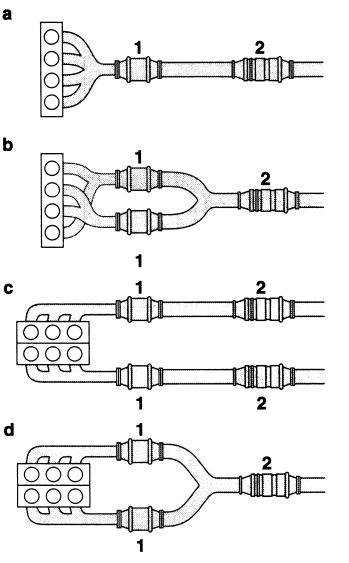
In the case of the NO_x storage catalytic converter, the favorable temperature range is lower; it reaches maximum storage capacity at 300 to 400 °C. The reason for the lower temperature range is that the maximum storage capacity is reduced at higher temperatures. At temperatures higher than 500 to 550 °C the barium compound is no longer stable, which means that nitrogen oxides can no longer be stored.

Operating temperatures of 800 °C to 1,000 °C result in intensified thermal aging of the catalytic converter. This is caused by sintering of the precious metals and the substrate coating, which reduces the active surface. At temperatures above 1,000 °C, thermal aging increases considerably until the catalytic converter has no effect at all.

Catalytic-converter configurations
The required operating temperature of the three-way catalytic converter limits the installation options. Engine-proximate catalytic converters quickly reach operating temperature, but can be exposed to very high thermal loads.

Figure 2: Catalytic-converter configurations
a) Deployment of a close-coupled catalytic converter and an underfloor catalytic converter.
b) 4-into-2 exhaust manifold for power-optimized engine design: The positioning of the underfloor catalytic converter only after the second junction is unfavorable for the heating-up characteristics, which means that preferably two close-coupled catalytic converters are deployed.
c) Engine with more than one cylinder bank (V engine): The exhaust-gas system runs completely as twin-branch with an upstream and an underfloor catalytic converter for each.
d) Engine with more than one cylinder bank (V engine): Y-shaped junction in the underfloor area to form an overall exhaust-gas train with a joint underfloor catalytic converter for both banks.

1 Close-coupled catalytic converter,
2 Underfloor catalytic converter or coated particulate filter.



A widely used configuration of the three-way catalytic converter is the split arrangement with a close-coupled catalytic converter and an underfloor catalytic converter. The close-coupled catalytic converter is optimized with regard to high-temperature stability, the underfloor catalytic converter with regard to "low light off" (low activation temperature). Figure 2 shows different possible arrangements of the primary and underfloor catalytic converters. Owing to their lower maximum permissible operating temperature, NO_x storage catalytic converters are always installed in the underfloor area.

Catalytic-converter heating

Whereas a catalytic converter at operating temperature reaches very high conversion rates of almost 100 %, considerably greater amounts of pollutants are emitted in the cold start phase and in the warm-up phase. The HC and CO emissions are particularly high when the engine is cold because fuel condenses onto the cold cylinder walls, leaves the combustion chamber unburned and cannot be converted by the cold catalytic converter.

It is therefore important to minimize untreated emissions during the warm-up phase before the catalytic converter lights off (operating temperature reached). Measures are also required to bring the catalytic converter up to operating temperature quickly. The required heat is provided by raising the exhaust-gas temperature and increasing the exhaust-gas mass flow. This is made possible by the following measures.

Ignition-angle adjustment

The main measure for increasing the exhaust-gas heat flow is ignition-angle adjustment. Combustion is initiated as late as possible and takes place during the expansion phase. The exhaust gas has a relatively high temperature at the end of the expansion phase. Retarded combustion has an unfavorable effect on engine efficiency.

Idle speed increase

A supporting measure is to raise the idle speed and thereby increase the exhaust-gas mass flow. The increased engine speed permits a greater ignition-angle retardation. However, in order to ensure reliable ignition, the ignition angles are limited to roughly 10 to 15 °KW after TDC. The heat output limited in this way is not always enough to achieve the current emission limits.

Exhaust-camshaft adjustment

A further contribution to increasing the heat flow can be achieved, if necessary, with exhaust-camshaft adjustment. The process of the exhaust valves opening as early as possible interrupts the retarded combustion early and the mechanical work generated is thus reduced further. The quantity of energy not converted into mechanical work is available as a quantity of heat in the exhaust gas.

Homogeneous split

With gasoline direct injection, there is, in principle, the possibility of multiple injection. This enables the catalytic converter to be quickly heated up to operating temperature without the need for additional components. The "homogeneous-split" measure involves initially creating a homogeneous, lean basic mixture by means of injection during the induction stroke. Subsequent injection during the compression stroke generates a stratified charge cloud. This facilitates retarded moments of ignition and results in high exhaust-gas heat flows. The exhaust-gas heat flows that can be achieved are comparable with those of a secondary-air injection.

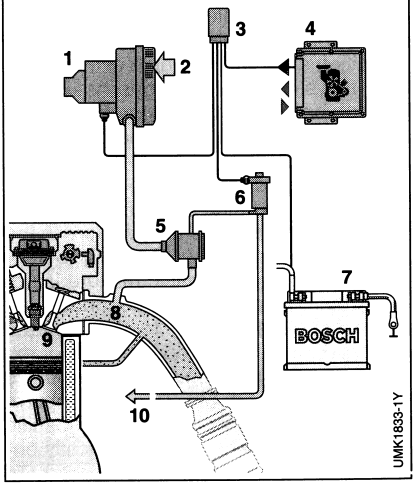
Secondary-air injection

Thermal afterburning of unburnt fuel constituents increases the temperature in the exhaust system. A rich (λ = 0.9) to very rich (λ = 0.6) air/fuel basic mixture is adjusted for this purpose. A secondary-air pump supplies oxygen to the exhaust system (Figure 3) to produce a leaner exhaust-gas composition. Where the basic mixture is very rich (λ = 0.6), the unburnt fuel constituents oxidize above a specific temperature threshold exothermically before entering the catalytic converter. To

achieve this temperature, it is necessary on the one hand to raise the temperature level with retarded ignition angles and on the other hand to introduce the secondary air as closely as possible to the exhaust valves. The exothermic reaction in the exhaust system increases the heat flow to the catalytic converter and therefore shortens its heating period. The HC and CO emissions are for the most part reduced before entering the catalytic converter.

Where the basic mixture is less rich (λ = 0.9), there is no significant reaction ahead of the catalytic converter. The unburnt fuel constituents oxidize in the catalytic converter and heat it up from the inside. For this purpose, however, the end face of the catalytic converter must first be brought up to a level above the light-off temperature by means of conventional measures (e.g. ignition-angle adjustment).

Figure 3: Secondary-air system
1 Secondary-air pump,
2 Inducted air,
3 Relay,
4 Engine control unit,
5 Secondary-air valve,
6 Control valve,
7 Battery,
8 Point of introduction into exhaust pipe,
9 Exhaust valve,
10 To intake-manifold connection.



As a rule, a less rich basic mixture is adjusted, because, if the basic mixture is very rich, the exothermic reaction ahead of the catalytic converter occurs reliably only under stable boundary conditions.

Secondary-air injection is performed with an electric secondary-air pump, which is switched by means of a relay on account of the high power demand. Since the secondary-air valve prevents backflow of exhaust gases into the pump, it must remain closed when the pump is deactivated. Either it is a passive non-return valve or it is actuated by purely electrical or (as shown in Figure 3) pneumatic means with an electrically actuated control valve. When the control valve is actuated, the secondary-air valve opens in response to the intake-manifold vacuum. The secondary-air system is coordinated by the engine ECU.

Alternative concepts for active heating
As a supplement in special cases, an electrically heated catalytic converter is used for quick heating up of the catalytic converter. This has been used previously in individual small-series projects.

λ control

To insure that the conversion rates of the three-way catalytic converter are as high as possible for the pollutant components HC, CO and NO_x, the reaction components must be present in the stoichiometric ratio. This requires a mixture composition of $\lambda = 1.0$, i.e. the stoichiometric air/fuel ratio must be adhered to very precisely.

Mixture formation must be followed up in a control loop, because sufficient accuracy cannot be achieved solely by controlling the metering of the fuel. When the λ control loop is used, deviations from a specific air/fuel ratio can be detected and corrected through the quantity of fuel injected. The residual-oxygen content in the exhaust gas, which is measured with λ oxygen sensors (see Two-step λ oxygen sensor and Broad-band λ oxygen sensor), serves as the measure for the composition of the air/fuel mixture.

Two-step λ control

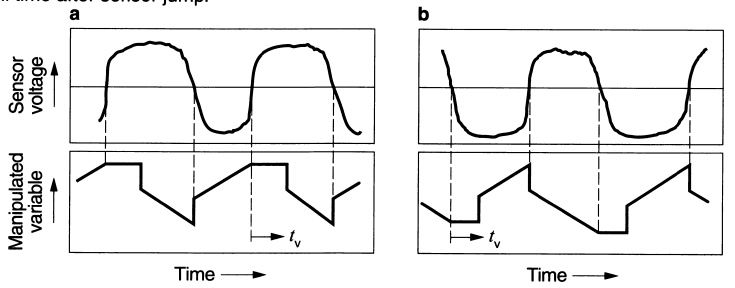
Two-step λ control adjusts the mixture to $\lambda = 1$. A manipulated variable, composed of the voltage jump and the ramp, changes its direction of control for each voltage jump of the two-step λ sensor. This indicates a change from rich to lean or from lean to rich (Figure 4). The typical amplitude of this manipulated variable has been set in the range of 2 to 3 %. The result is a limited controller dynamic, which is predominantly determined by the sum of the response times (pre-storage of fuel in the intake manifold, four-stroke principle of the spark-ignition (SI) engine, and gas travel time).

The typical shift of the oxygen zero transition point (theoretically at $\lambda = 1.0$) and therefore the jump of the λ sensor, caused by the variation in exhaust-gas composition, can be control-compensated by shaping the manipulated variable's characteristic curve asymmetrically (rich or lean shift). Here, the preferred method is to hold the ramp value at the voltage jump for a controlled dwell time t_V after the sensor is subjected to a voltage jump. A rich or lean shift is performed depending on the operating point. When a shift is made towards "rich", the manipulated variable holds at the rich position for a dwell time t_V although the sensor signal has already jumped towards "rich" (Figure 4a). Jump and ramp of the manipulated variable do not also move towards "lean" until the dwell time has elapsed. If the sensor signal then jumps towards "lean", the manipulated variable directly counteracts this (with jump and ramp) without dwelling on the lean position. The dwell time t_V is the result of a pre-control coefficient, which is corrected by a proportion from the two-sensor λ control system (post cat control).

The behavior is reversed in a shift to "lean". If the sensor signal indicates a lean mixture, the manipulated variable holds at the lean position for the dwell time t_V (Figure 4b) and then adjusts towards "rich". A

Figure 4: Manipulated-variable curve of a two-step control with a λ sensor upstream of the catalytic converter and λ shift (delay time t_v) due to pre-control proportion and post cat control

a) Shift to "rich", b) Shift to "lean".
t_v Dwell time after sensor jump.



sensor signal jump from "lean" to "rich" is, however, immediately counteracted.

Continuous λ control

The defined dynamic response of a two-step control can only be improved if the deviation from $\lambda = 1$ can actually be measured. The broad-band λ sensor can be used to achieve continuous-action control at $\lambda = 1$ with a stationary, very low amplitude in conjunction with high dynamic response. The control parameters are calculated and adapted as a function of the engine's operating points. Above all, with this type of λ control, compensation for the unavoidable offset of the stationary and non-stationary pilot control is far quicker.

The broad-band λ oxygen sensor also enables adjustment to mixture compositions that deviate from $\lambda = 1$. Thus, controlled enrichment ($\lambda < 1$), e.g. for component protection, can be effected. The broad-band λ sensor also enables controlled leaning ($\lambda > 1$), e.g. for a leaner warm-up during catalytic-converter heating.

Two-sensor λ control

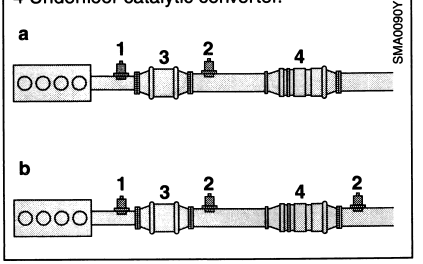
When it is situated upstream of the catalytic converter, the λ sensor is heavily stressed by high temperatures and untreated exhaust gas, and this leads to limitations in λ sensor accuracy. The voltage-jump point of a two-step sensor or the characteristic curve of a broad-band sensor can shift, for example, as a result of changed exhaust-gas compositions. A λ sensor downstream of the catalytic converter is subjected to these influences to a much lesser extent. However, λ control on its own with the sensor downstream of the catalytic converter demonstrates disadvantages in dynamic response on account of the gas travel times, and responds to mixture changes more slowly.

Greater accuracy is achieved with two-sensor control. Here, a slower correction control loop is superimposed on the two-step or continuous-action λ control described by means of an additional two-step λ sensor (Figure 5a). The voltage of the two-step sensor downstream of the catalytic converter is compared with a setpoint value (e.g. 600 mV) for this purpose. On this basis the control evaluates the deviations from the setpoint value and additionally alters over a dwell time t_v additively the controlled rich or lean shift of the first control loop of a two-step control or the setpoint value of a continuous-action control.

Three-sensor λ control

There are SULEV concepts (Super Ultra-Low-Emission Vehicle) with a third sensor downstream of the underfloor catalytic converter. The two-sensor control system (single cascade) has been supplemented here by an extremely slow closed-loop control system using a third sensor located downstream of the underfloor catalytic converter (Figure 5b). The aim of development is, however, the two-sensor concept.

Figure 5: Installation locations of λ sensors
a) Two-sensor control,
b) Three-sensor control.
1 Two-step or broad-band λ sensor,
2 Two-step λ sensor,
3 Close-coupled catalytic converter,
4 Underfloor catalytic converter.



Particulate filter for gasoline engines

Function

The function of the particulate filter in the gasoline engine with gasoline direct injection is to adhere to the limit values for particulate matter and particulate number laid down in the exhaust-emissions standard Euro 6c or Euro 6d(-temp) (see also Particulate filter for diesel engines). An additional catalytic coating of the particulate filter serves to increase the catalytic-converter volume available for the oxidation of CO and HC and the reduction of NO_x. A catalytically coated particulate filter can thus replace a three-way catalytic converter.

Design

The particulate filter for gasoline engines is designed in the same way as the particulate filter in diesel engines (see Particulate filter for diesel engines). For gasoline applications the ceramic is currently made of cordierite.

Regeneration of the particulate filter

Soot is oxidized by oxygen only at temperatures above approximately 550 to 600 °C. In gasoline applications of the particulate filter the temperatures reached in normal operation and the requirement of additional engine measures to increase the exhaust-gas temperature are determined by the filter installation position. Usually the filter is positioned downstream of the three-way catalytic converter. When installed directly downstream of the three-way catalytic converter temperatures of 550 to 600 °C can be achieved in part-load operation without additional measures, while this is only achieved with an installation position in the underfloor area in high-load operation. In terms of regeneration performance an installation position directly downstream of the three-way catalytic converter therefore exhibits clear advantages, but the space required for this must be available near the engine. Integrating the filter function into the three-way catalytic converter for the four-way catalytic converter (FWC) does exhibit clear advantages with regard to the regeneration performance and near-engine installation area, but up to now has only been implemented in individual applications due to the highest thermal load at full load.

The most important engine measure to increase the exhaust-gas temperature is, as with heating the catalytic converter, ignition-angle retardation, which simultaneously brings about de-throttling and an increase in the exhaust-gas mass flow and in the case of gasoline direct injection can be combined with multiple injection ("homogeneous split"). Another measures is the alternate rich/lean adjustment of individual cylinders while adhering to an exhaust-gas λ of 1. Unburned exhaust-gas constituents from the rich-burn cylinders are exothermally converted here with O_2 from the lean-burn cylinders in the three-way catalytic converter.

In contrast a diesel engine, in a gasoline engine operated with exhaust-gas λ of 1 there is no O_2 available for soot oxidation in normal operation. After the required temperature is reached, the soot can therefore be burned off initially in an overrun phase. Alternatively, it is possible to achieve soot burn-off by setting a lean exhaust-gas λ of up to approximately 1.1, which is however to be avoided because of the increased NO_x emissions involved.

Saturation detection

The strategies used to determine the amount of soot stored in the particulate filter are comparable with the method used on diesel engines (see Particulate filter for diesel engines).

Motor sport

Requirements

The basic requirements regarding components in motor-sport vehicles are different from those in production vehicles. A racing car requires maximum performance for a few hours with the lowest possible weight, whereas durability and low costs are the desired attributes in a production car. Furthermore, the components in a racing car are exposed to greater loads from heat, dust, moisture, and vibrations than those in a production car.

The Bosch Motorsport Division develops, produces and sells electronic systems and components for this special field of application. Both modified mass-produced parts and pure motor-sport developments, which often have no equivalent in series production, are offered.

Components

Modified mass-produced parts

Reducing weight

Technicians are constantly on the lookout for ways of reducing the weight of production components for racing applications. The reduced service life generally associated with this is tolerated.

By using ultralight special materials and specific special constructions, it is possible to reduce the weight of many components – and all this with higher power output. As the motto goes: Remove as much material as possible, keep as little material as is necessary. A quotation that goes around in design-engineer circles puts it in a nutshell: "The perfect racing car crosses the finishing line in first place and then falls apart into its component parts".

Increasing performance

There is a huge range of measures to increase the performance of production components. Alternators for instance have additional windings and thereby produce more energy. Their rotors are fine-balanced and are those less sensi-

Figure 1: Starting grid for a Deutsche Tourenwagen Masters (DTM) race (photo: Bosch)

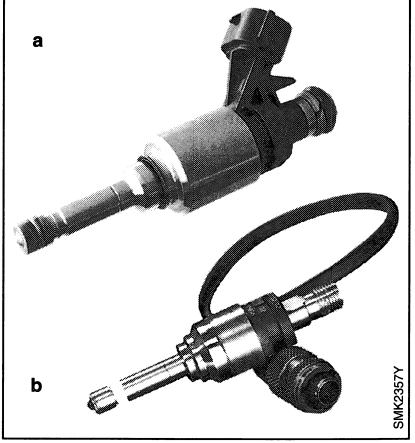
tive to high rotational speeds. Temperature sensors respond much more quickly once the metal that normally encloses the measuring capsule has been removed. Changes to fuel pumps and fuel injectors are frequently made with increased throughflow and increased pressure in mind.

The sometimes significantly higher specific engine performance of racing cars makes these measures absolutely necessary. The use of better ferromagnetic materials generate higher magnetic forces in the coils of the fuel injectors. This makes possible greater needle strokes, which in turn permit a higher fuel flow.

The modifications to ignition coils go even further. Here only the original shell remains the same, the technology inside is all new. The coils are made with fewer windings from thicker wire, the coil cores are made of higher-quality material. The higher spark energy required for direction-injection highly supercharged racing engines can be guaranteed in this way. The shorter dwell periods increase the upper speed limit to 15,000 rpm.

Figure 2: High-pressure fuel injectors in motor sport

a) Component with production connector on the housing.
b) Component with cable and motor-sport connector.



Replacing connectors

Production connectors made of plastic are not designed for frequent opening and closing and have only limited resistance to vibrations and moisture. They are therefore often removed from motor-sport components and replaced by cables with motor-sport connectors (Figure 2). Their metal housings have a bayonet catch for reliable and rapid opening and closing. They are dust-proof and watertight and, because they are no longer rigidly connected to the component, can be installed in locations that are less exposed to vibration. Particularly for sensors, ignition coils and fuel injectors this modification makes installation and removal easier and increases operational reliability by minimizing primarily vibration-induced damage. These connectors are standard in motor racing and are compatible with virtually all wiring harnesses and diagnostic tools.

Dissipating heat

High-performance electronic systems generate a lot of heat which can jeopardize temperature-sensitive components. This heat is dissipated by cooling fins to the housing walls and the use of thermal-conduction paste in the housings. These housings are often black to improve heat dissipation.

Regular maintenance

Unlike in production-vehicle manufacturing, where starters and alternators as a rule go through the entire life of the vehicle without maintenance and repairs, these components are overhauled in motor racing after a certain period of operation. Bosch offers its customers this service for all motor-sport components. The components are dismantled, cleaned and provided with new wearing parts at the factory. They are then restored to the same operational reliability as new parts.

ECUs

The fundamental difference from a production ECU, where the manufacturer must guarantee tuning protection measures, is the free access to the software for all motor-sport systems. In this way, the teams can tune each car individually to racetrack, weather and driver. The computer programs required for tuning on the laptop computer are developed by Bosch itself. Furthermore, programs for analyzing recorded vehicle data and for simulating different vehicle tunings are offered.

Series-based ECUs
Series-based developments with sheet-metal housings and plastic connectors constitute the entry level into motor-sport ECUs. The range of functions is less complex than in high-end ECUs and can therefore also be easily handled by smaller teams. These systems are frequently used in Formula 3, in marque cups and in near-series motor-racing classes, where many GT3 class customer sports cars are used.

High-end ECUs
Top-end models among the ECUs have all-aluminum milled housings and motor-sport connectors (Figure 3). The housings are powder-coated to protect the surface. The number of connector pins and hence of inputs and outputs is high, and the number of controllable cylinders is frequently high too. Some systems have customer code areas, in which customers can store their own software and link with the basic software. Communication channels to CAN and LIN bus systems as well as telemetry connections and internal data memories are standard.

Working with these systems calls for a high level of specialized knowledge, which as a rule can only be brought to bear by works teams. Typical areas of application for these ECUs are the World Rally Championship (WRC), the World Endurance Championship (WEC), and the Deutsche Tourenwagen Masters (DTM).

Pure motor-sport developments

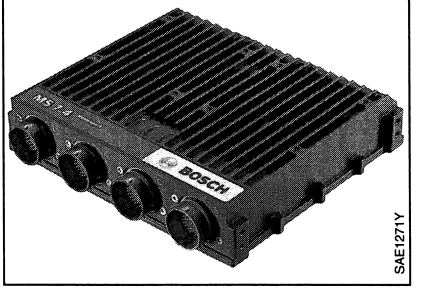
As well as the high-end ECUs, further pure motor-sport developments are featured in the Bosch Motorsport portfolio. These include, for example, lap-trigger systems with which the times of complete laps or lap segments are measured and telemetry systems for data transfer from the racing car to the box. Driver displays replace all the instruments in the dashboard, are freely programmable, and have switchable display pages. They are frequently equipped with data memories.

High-pressure fuel-injection systems and their valve injectors often require a significantly higher operating voltage than the conventional vehicle electrical system can provide. This gap is filled by the power supply unit which is used in many Formula 1 cars and delivers voltages of up to 90 V.

Safety and comfort
Electronic systems from Bosch however not only deliver better acceleration, but also better deceleration and increased safety and comfort in racing cars.

The Bosch motor-sport ABS systems come with different control characteristics which the driver can select by means of a rotary switch in the dashboard. Here too the ECU software is open and can be changed where necessary. Pre-parameterized plug-and-play versions are offered for certain cup cars. The experts from Bosch lend the customers their sup-

Figure 3: High-end ECU with milled housing and motor-sport connectors



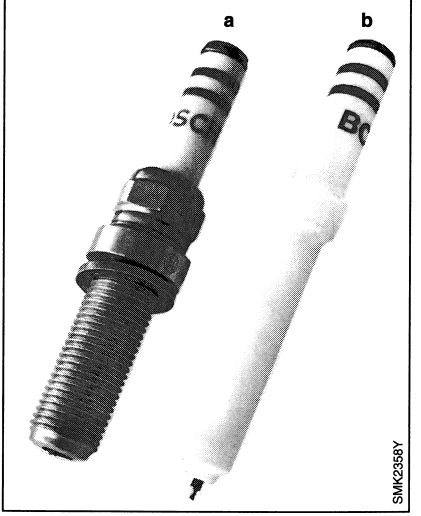
port during individual tuning of the ABS systems on the test bench or on the racetrack.

A radar-based accident-avoidance system uses visual signals to warn against passing cars; these systems are particularly helpful when it is raining or at night.

The Bosch Power-Box replaces all the relays and fuses in a racing car and enables the relay switching points to be changed on a laptop computer without a mechanic having to undo a screw or exchange a component. The laptop computer here also replaces the tool box.

In the case of systematic troubleshooting it is very helpful to be able to measure specifically the signal flow of individual cables in the wiring harness. Here a specifically developed breakout panel, which is switched between the wiring harness and the ECU, provides sterling service.

Figure 4: Spark-plug variants for racing use
a) Racing spark plug with 8 mm thread, toothing for special tool and platinum electrode.
b) Stripped-back spark plug for subsequent processing by the motor-racing customer.



Customer-specific extra-small series

Bosch hand-makes in small numbers motor-sport wiring harnesses which network the systems in a racing car with each other. They have the customary motor-sport connectors and are delivered to the customer after exhaustive functional tests.

Motor-sport spark plugs only have the basic idea in common with the mass-produced components. Electrodes made of pure platinum permit operation at significantly higher temperatures. Thread diameters reduced to 8 mm allow the engine developers plenty of freedom in fashioning the combustion chambers and the cooling ducts (Figure 4). Thread length, width across flats and other sizes are designed by arrangement with the customer. This results in hand-made sample spark plugs which are specially customized to the existing engine concept.

As well as the ready-to-install spark plugs, Bosch offers its motor-racing customers the separate ceramic core of the spark plug for subsequent processing. The customer provides this "stripped-back spark plug" with their own shell, the design of which is among the best-kept secrets of the design engineers.

New directions

Mirroring developments in production-vehicle manufacturing, hybrid systems are increasingly being used in racing cars. Electric motors work in coordination with gasoline or diesel engines; recovered energy is buffered in electrical or mechanical accumulators.

Some motor-racing classes are already fully reliant on the purely electric drivetrain, as is demonstrated for example by Formula E cars, which race on city-center street circuits around the world.

In these new directions too the long tradition of exchanging experience and technology between series production and motor sport is being continued by Bosch.

LPG operation

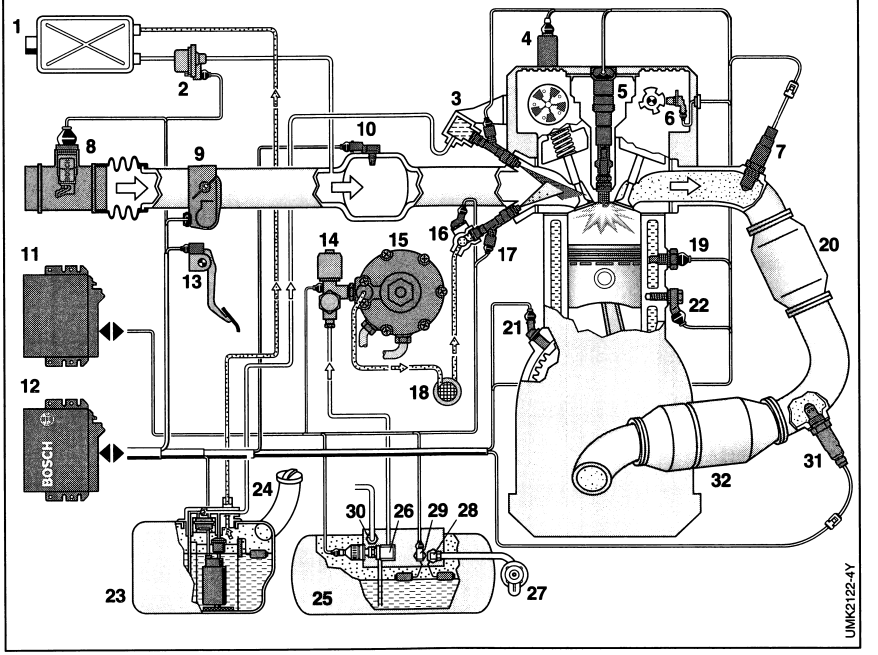
Applications

Liquified petroleum gas (LPG) is obtained during the extraction of crude oil and natural gas and during different refinery processes. It is made up primarily of propane and butane (see Liquid petroleum gas).

As the conventional fuels gasoline and diesel have become more expensive in the last few years, the cheaper liquified petroleum gas – not least due to regional tax reductions (e.g. for Germany, Italy, Poland) – is becoming increasingly more widespread. Liquified petroleum gas is widely used in Europe and in South Korea, Australia, and China. At present (as at No-

Figure 1: Bivalent gasoline-engine operation with gaseous LPG injection

1 Activated charcoal canister, 2 Canister-purge valve, 3 Gasoline rail with gasoline injectors,
4 Camshaft adjuster, 5 Ignition coil with spark plug, 6 Camshaft sensor,
7 λ sensor upstream of primary catalytic converter, 8 Hot-film air-mass meter,
9 Electronically actuated throttle valve (ETC), 10 Intake-manifold pressure sensor,
11 LPG ECU, 12 Engine ECU, 13 Accelerator-pedal module,
14 LPG shutoff valve, 15 Evaporator with pressure regulator,
16 Combined pressure and temperature sensor, 17 LPG rail with gas injectors,
18 LPG filter, 19 Engine-temperature sensor, 20 Primary catalytic converter,
21 Crankshaft speed sensor, 22 Knock sensor,
23 Gasoline tank with integrated electric fuel pump, 24 Gasoline filler neck,
25 LPG tank (steel tank), 26 Electromagnetic shutoff valve,
27 LPG filler neck, 28 80% filler stop valve, 29 LPG level indicator,
30 Pressure-relief valve, 31 λ sensor downstream of primary catalytic converter,
32 Main catalytic converter.



vember 2016), there is estimated to be a population of more than 26 million LPG-fueled vehicles worldwide [1]. Up-to-date information on the number of LPG-fueled vehicles and the LPG filling-station network in Germany (currently around 75,350 filling stations worldwide, around 7,050 filling stations in Germany) can be found on the Internet (see [2] and [3]).

The additional volumetric consumption of LPG-fueled vehicles is, depending on engine concept, engine management and gas system 20–30 %, compared with gasoline-fueled vehicles. In spite of this additional costs, the running costs of LPG vehicles are cheaper than those of gasoline and diesel vehicles.

LPG systems are distinguished by their mixture formation with gaseous injection or liquid LPG injection and by the ECU concepts with one common or two separate engine ECUs (master/slave concept). Systems with venturi mixers (operating according to the carburetor principle) can only still be used in older vehicles or in markets with laxer exhaust-emission regulations.

Vehicles with spark-ignition engines are virtually exclusively converted to bifuel LPG/gasoline operation. Most LPG systems are installed as a retrofit. In the meantime, however, vehicle manufacturers are also offering models with bifuel LPG/gasoline operation off the assembly line. In these vehicles, the LPG tank – as in the retrofit concepts – is installed for the most concepts in the spare-wheel compartment. The luggage-compartment load volume remains unchanged with this form of installation. The spare wheel is simply omitted and replaced by a blow-out repair kit (Tire-Fit).

Design

Storage of LPG

LPG can liquify at ambient temperatures of –20 to +40 °C under pressure (3–15 bar), depending on the propane-to-butane mixture ratio. It is stored in a liquid state in steel tanks. The filling system of the LPG tanks incorporates a fill-volume restrictor which ensures that the fill volume does not exceed 80 % of the liquid phase of the maximum tank volume. This ensures that the internal tank pressure does not rise to unacceptable levels when the tank is subjected to heat. The tanks are subject (also in some non-European countries) to European safety regulations governing their properties, equipment specification and installation – currently ECE-R67-01 [4].

The safety devices of an LPG tank comprise a pressure-relief valve, an electromagnetic shutoff valve with integrated flow limiter (pipe-fracture protection), and optionally a thermal safety valve which allows the gas to escape under controlled conditions in the event of a fire.

Fuel delivery

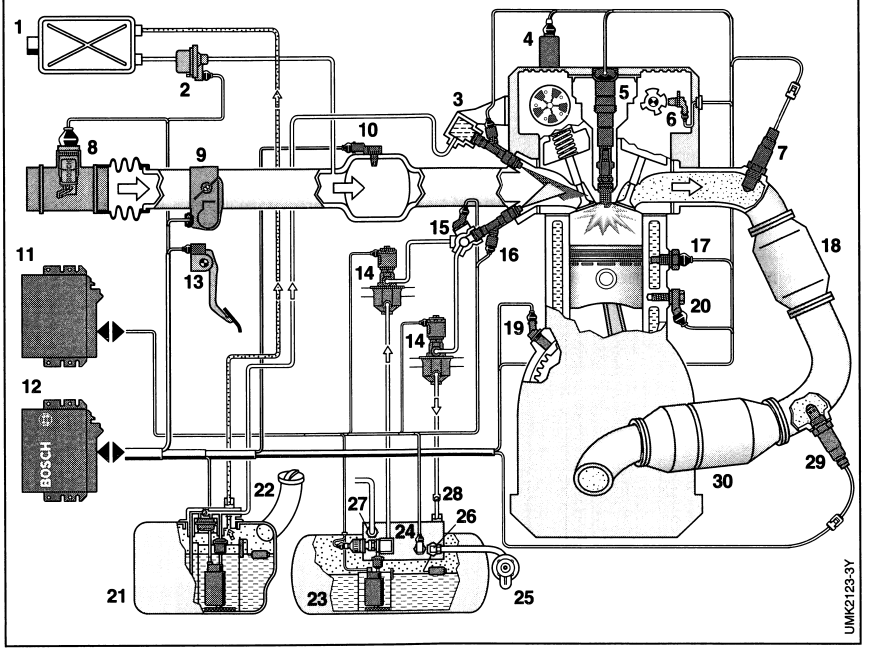
The LPG is delivered from the tank to the engine in systems with gaseous injection by the internal tank pressure (LPG vapor pressure). Because this pressure drops as ambient temperature decreases, the operation of such systems is partly limited with regard to their suitability for winter operation.

In systems with liquid injection, the LPG is delivered by an LPG fuel-supply pump installed in the LPG tank which generates a delivery pressure in addition to the internal tank pressure. In terms of their suitability for winter operation, these systems are comparable with gasoline injection systems.

Components
The following components are integrated into LPG vehicles (Figures 1 and 2):
- LPG filler neck with integrated non-return valve,
- Steel tank (cylindrical or toroidal),
- Pressure-relief valve (27 bar),
- LPG level sensor,
- Valve with integrated 80 % tank filler stop,
- Electromagnetic shutoff valve (in the tank),
- LPG shutoff valve (in the engine compartment),
- Evaporator with pressure regulator (only for systems with gaseous LPG injection),
- Pressure regulator (optional in systems with liquid LPG injection),
- LPG fuel rail with gas injectors,
- Pressure and temperature sensors (on the rail),
- LPG filter,
- If necessary, a slave ECU,
- Demand-controlled LPG supply pump (only in systems with liquid LPG injection).

Figure 2: Bivalent gasoline-engine operation with liquid LPG fuel injection into the intake manifold

1 Activated charcoal canister, 2 Canister-purge valve, 3 Gasoline rail with gasoline injectors,
4 Camshaft adjuster, 5 Ignition coil with spark plug, 6 Camshaft sensor,
7 λ sensor upstream of primary catalytic converter, 8 Hot-film air-mass meter,
9 Electronically actuated throttle valve (ETC), 10 Intake-manifold pressure sensor,
11 LPG ECU, 12 Engine ECU, 13 Accelerator-pedal module,
14 LPG shutoff valves for supply and return, 15 Combined pressure and temperature sensor,
16 LPG fuel rail with gas injectors, 17 Engine-temperature sensor,
18 Primary catalytic converter, 19 Crankshaft-speed sensor, 20 Knock sensor,
21 Gasoline tank with integrated electric fuel pump, 22 Gasoline filler neck,
23 LPG tank with integrated electric LPG pump, 24 Electromagnetic shutoff valve,
25 LPG filler neck, 26 80 % filler stop valve with integrated LPG level sensor,
27 Pressure-relief valve, 28 Non-return valve,
29 λ sensor downstream of primary catalytic converter, 30 Main catalytic converter.



LPG systems

Systems with gaseous LPG injection

Operating principle

The LPG is delivered in liquid state from the tank via the electromagnetic shutoff valve to the engine compartment (Figure 1), where an additional shutoff valve is installed. After this, the LPG is converted into the gaseous state with pressures of 0.5 to 3.5 bar in the evaporator and the pressure regulator. The evaporator is heated by the engine coolant in order to compensate for the cooling encountered during the phase transition. The gaseous LPG is directed from the evaporator through a pressure-resistant flexible hose to the LPG rail and from here injected by the gas injectors into the intake manifold.

The air-intake section of the basic engine remains unchanged except for the connection of the LPG gas injectors to the intake manifold.

Mixture formation

Mixture formation is effected in the same manner as conventional multipoint gasoline injection. The LPG rail supplies the individual gas injectors with gaseous LPG. Engine management determines the required injected fuel quantity, which is corrected by means of λ control. The injected fuel quantity is additionally corrected by means of the pressure and the temperature of the fuel in the LPG rail.

LPG injection displaces some of the inducted air volume in the intake manifold, which results in charge losses in non-turbocharged engines. Engine performance reduced by a factor of 2–4% compared with gasoline operation can thus be expected for LPG operation.

Systems with LPG liquid injection into the intake manifold

Operating principle

In systems with LPG liquid injection into the intake manifold (Figure 2) the LPG ECU regulates the system pressure via a fuel pump installed in the LPG tank.

Some systems may also need a pressure regulator to regulate the injection pressure. Because of the tendency for vapor bubbles to form in the fuel system, systems with LPG liquid injection must be fitted with a tank return line in order if necessary to expel vapor bubbles that occur.

Mixture formation

As with conventional multipoint gasoline injection, the LPG is injected into the intake manifold in front of the engine intake valves. The evaporation that occurs as the LPG is injected (according to the Joule-Thomson effect) lowers the temperature of the inducted air/gas mixture, which has a positive impact on the engine's volumetric efficiency and performance figures.

Advantages of LPG liquid injection

LPG liquid injection offers the following advantages:

- Gas operation is also possible at extremely low temperatures.
- Monofuel gas operation is possible.
- Engine start in LPG mode is also possible in bifuel systems.
- A better cylinder charge is obtained thanks to the cooling effect in the intake manifold and combustion chamber, thus producing identical engine-performance figures to gasoline without further measures. Engines with LPG liquid injection can even deliver slightly higher performance figures when compared with gasoline operation.

Systems with direct LPG liquid injection into the combustion chamber

LPG behaves very similar to gasoline at higher system pressures. Therefore LPG direct injection (LPG-DI) into the combustion chamber is also possible. LPG direct injection offers some significant advantages:

- A better exhaust-gas composition is obtained (virtually no particulate emissions) with comparable engine-performance figures.
- Very good over all thermodynamic behavior.

LPG direct liquid injection into the combustion chamber is possible in all modern gasoline engines with direct injection. Such systems are currently still in the development phase (see e.g. [5]).

Exhaust emissions

With regard to the limited pollutant components HC, CO and NO_x, the use of LPG as fuel in modern spark-ignition engines offers some advantages over gasoline operation. The situation when it comes to CO_2 emissions is even better. Here, LPG vehicles can achieve a reduction of approximately 10 %. This reduction is founded on the lower carbon content in LPG. Furthermore, when compared with diesel engines, there are virtually no particulates. LPG drives also have advantages when it comes to other exhaust-gas components which are not limited by law but which are harmful to health. Thus, for instance, LPG drives emit significantly lower amounts of aromatic hydrocarbons (e.g. benzene).

Unlike engines fueled by CNG (Compressed Natural Gas), engines fueled by LPG are not subject to any particular demands on the exhaust-gas treatment system of the existing gasoline concept. The essential exhaust-gas components of LPG operation are comparable with those of gasoline operation.

Engine management for LPG systems

Both double-ECU concepts (one ECU for gasoline operation and one for LPG operation) and single-ECU concepts (gasoline and gas functions integrated into a single ECU) are used. Because the vehicles are bifuel, the driver can select between gasoline and LPG operation using a switch.

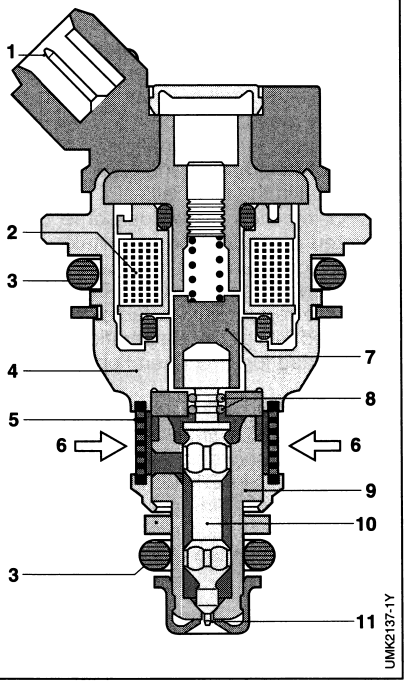
The level sensor installed in the LPG tank signals the current LPG tank level to the engine-management system. A combined low-pressure and temperature sensor mounted on the gas rail enables the engine-management system to correct the injection timing in such a way that a stoichiometric mixture can be pilot-controlled in the intake manifold even if the gas density fluctuates. The engine-management systems uses the values from the λ sensor and different adaptation algorithms to adapt the mixture formation for different gas qualities. Even functions for switching between gas and gasoline operation are integrated.

The other sensors and actuators of the engine-management system are for the most part identical to those in a gasoline engine.

Systems with LPG liquid injection into the combustion chamber require, in addition to the functions for mixture formation and gas-quality adaptation, functions for controlling the LPG fuel pump.

Figure 3: Electromagnetic liquid LPG injector

1 Electrical connection,
2 Field winding,
3 O-rings,
4 Valve housing,
5 Filter strainer,
6 LPG supply,
7 Armature,
8 O-rings,
9 Valve body,
10 Valve needle,
11 Injection nozzle.



Components

LPG gaseous injector

Function and requirements

When LPG is injected in the gaseous state of aggregation, the given energy density and the low gas pressures resulting from the state of aggregation of approx. 0.5 to 3.5 bar dictate that the following parameters must be taken into account in the design of LPG injection valves:

- Configuration of the injection valves for large gas volumes,
- Adaptation of the sealing mechanisms and materials to the gaseous medium,
- Consideration of potentially aggressive fuel contaminants in the choice of material.

Design and operating principle

Essentially, the design and function of a gas injector configured for operation with gaseous LPG are very similar to those of a CNG injector (see Engines fueled by natural gas).

LPG liquid injector

Function and requirements

In systems where the LPG is injected in the liquid state of aggregation the LPG must be kept in the liquid state in the entire pipe and rail system by means of pressure build-up by a pump installed in the LPG tank. Only in this state can it be injected.

LPG injectors must satisfy the following requirements:

- Operation at variable pressures of 2 to 25 bar.
- Absolute tightness in the case of systemic pressure variability (e.g. on shutdown of gas operation and spontaneously occurring pressure peaks due to the LPG evaporating in the rail system).
- Possibility of a flushing process to eliminate vapor bubbles that occur (e.g. during hot start or switching from gasoline to LPG operation with the engine at operating temperature).

Design and operating principle

LPG injectors (Figure 3) are, in response to the above-mentioned requirements, designed according to the bottom-feed principle. The injector is designed in such a way that the fitting for the fuel inlet into the injector housing is as close as possible to the injector outlet. This layout enables the valve to be easily flushed with fresh fuel in order to quickly remove possible vapor bubbles.

The fuel injectors normally used in gasoline engines, on the other hand, operate for the most part according to the top-feed principle. Because of the supply of fresh fuel located at the top end of the injector housing, the flush process for the purpose of removing vapor bubbles in the injector would be very diffecult or not possible at all. Under such conditions it would not be possible to provide a correct metering of the fuel quantity and thus optimum combustion, which is why the top-feed principle is not suitable for use in liquid LPG injection.

Installation position

The position of the LPG injectors in the intake manifold is subject to special design requirements according to the generation of cold by the evaporation of the LPG during the injection process (as a result of the Joule-Thompson effect). The liquid LPG evaporates abruptly on entry into the intake manifold and extracts the required energy in the form of heat out of the fresh air in the intake manifold. This specific-point cold in the intake manifold can cause unwelcome operating phenomena (e.g. misfires), which can be put down among other things to the formation of ice crystals.

LPG filler neck

Neck variants

The filler neck is not internationally standardized. For this reason, different fueling adapters are required from country to country. Four different types of LPG fuel nozzle are used in filling stations. Different filler necks are therefore needed depending on the field of application. The following necks are widely used:

- ACME neck (named for the American trapezoidal-thread standardization) is used in Germany, Belgium, Ireland, England, Luxembourg, Scotland, Austria, China, Canada, USA, Australia, and Switzerland.
- DISH neck (named for its dish shape) is used in Denmark, France, Poland, Greece, Hungary, Italy, Austria, Portugal, the Czech Republic, Turkey, China, Korea, and Switzerland.
- Bayonet neck is used in the Netherlands, Belgium and the UK.
- Euronozzle is currently used in Spain only.

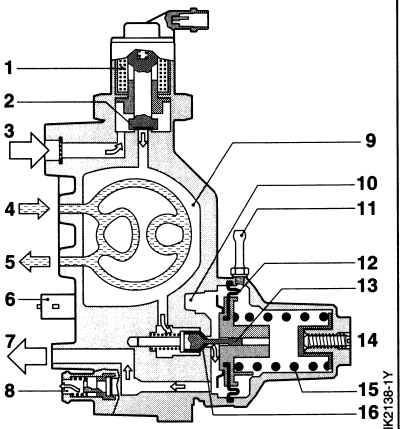
The Euronozzle is supposed to replace all the other necks on a Europewide basis in future. The EU Regulation for this standardized neck has not yet been adopted by the majority of EU member states.

Design

Each filler neck in the vehicle is designed in such a way that the LPG vehicle can be refueled at all LPG filling stations by using a suitable adapter.

Figure 4: Evaporator with pressure regulator

1 Solenoid coil of electromagnetic shutoff valve,
2 Seal,
3 Inlet for liquid LPG and filter,
4 Inlet for engine coolant,
5 Outlet for engine coolant,
6 Temperature sensor,
7 Outlet for gaseous LPG at constant pressure,
8 Safety pressure-relief valve,
9 Evaporation chamber,
10 Area for gaseous LPG,
11 Connection for intake-manifold vacuum,
12 Pressure diaphragm with mounting,
13 Control rod,
14 Adjusting screw,
15 Spring,
16 Control plunger and throttling orifice.



Evaporator for systems with gaseous LPG injection

Function

Evaporators are used in systems with gaseous LPG injection. The LPG is delivered to the evaporator by the vapor pressure in the gas tank via an electromagnetic shutoff valve.

The evaporator is one of the main components of every LPG injection system and performs two functions: Firstly, the liquid gas is converted from the liquid to the gaseous state in the first part of the evaporator. Secondly, the pressure of the gaseous LPG is then regulated to approx. 1 bar differential pressure in relation to the intake manifold. This pressure, with which the injection valves are supplied, must be kept constant through all the operating states and the varying vapor pressures of the different LPG qualities within minimum tolerances. The most widely used evaporators are those with single-stage pressure regulation. Figure 4 shows an evaporator with single-stage pressure regulation.

Operating principle
The LPG enters through the inlet connection into the evaporation area. This consists of a labyrinth that is heated by the vehicle's heating circuit to convert the incoming liquid gas into the gaseous state. Pressure of 4 to 15 bar are achieved here. The pressure varies depending on the temperature of the engine coolant and the composition of the LPG (mixture ratio of propane and butane).

The now gaseous medium passes through a transfer passage into the pressure-regulation area, which is closed by a diaphragm. A pre-tensioned spring is seated on the back of the diaphragm. The pre-tension can be adjusted with an adjusting screw in order to effect deviations from the factory output-pressure pre-setting.

The diaphragm is connected via a lever mechanism to a valve on the transfer-passage connection. When the pressure increases, the movement of the diaphragm actuates a control rod, which in turn alters the opening cross-section at the outlet of the pressure-regulation area. Thus the supply is reduced or closed off and no more gas can flow. When the gas pressure decreases in response to increased fuel reduction by the engine, the diaphragm is moved in the opposite direction. The valve is opened via the lever mechanism to allow gas to flow until again an equilibrium of forces acts on the diaphragm and the lever mechanism can close the transfer passage again.

An important characteristic of the quality of the evaporator is the lowest possible fluctuation of the operating pressure in different engine operating states together with the highest possible control precision over the operating time.

In certain countries, owing to contaminants in the LPG, the manufacturer recommends that the diaphragm be replaced after a certain operating time.

Intake-manifold vacuum support
In addition, the response of evaporator pressure regulation is optimized by connecting the intake-manifold pressure to the area above the diaphragm.

Operating values
The inlet pressure is 4 to 15 bar, depending on ambient temperature and gas quality. The output pressure is 0.5 to 3.5 bar, depending on the setting.

Electromagnetic shutoff valve and temperature sensor
The electromagnetic shutoff valve interrupts the gas supply when the engine is not in gas mode. The changeover from gasoline to LPG operation is initiated by the engine ECU as a function of the engine coolant temperature. The minimum coolant-temperature value for changeover is selected to ensure reliable evaporation of the LPG in the evaporator. A temperature sensor is integrated in the evaporator coolant circuit for this purpose. A cold start in gas mode is not possible with these systems.

References
[1] World LPG Gas Association, www.wlpga.org.
[2] www.gas-tankstellen.de.
[3] www.autogastanken.de.
[4] ECE-R67: Standard conditions on the:
I. Approval of specific equipment of motor vehicles using LPG in their propulsion system.
II. Approval of a vehicle fitted with specific equipment for the use of LPG in its propulsion system with regard to the installation of such equipment.
[5] K. Reif: Ottomotor-Management, 4th Ed., Verlag Springer Vieweg, 2015.

Engines fueled by natural gas

Applications

In view of the worldwide efforts to reduce CO_2 emissions (carbon dioxide), natural gas is becoming increasingly important as an alternative fuel. Compressed natural gas (CNG), not to be confused with liquified petroleum gas (LPG), consists primarily of methane. Liquified petroleum gas, on the other hand, consists of propane and butane (see fuels).

Emissions
Compared to gasoline, the combustion of CNG produces approximately 25% less CO_2. CNG thus generates the lowest local CO_2 emissions of the fossil fuels. Because of the lower CO_2 emissions, the petroleum tax on CNG is reduced in many countries.

Use of biogas
Using biogas (biomethane) can achieve a regenerative fuel supply that further reduces global greenhouse-gas emissions. Biomethane can be produced from plant remains, organic waste and liquid manure, or alternatively from electricity generated from renewable sources. Here, hydrogen is obtained by electrolysis from water and then reacts in a chemical reaction with CO_2 to produce methane. This methane gas obtained from renewable sources is called solar gas, wind gas or simply RE gas (renewable-energy gas).

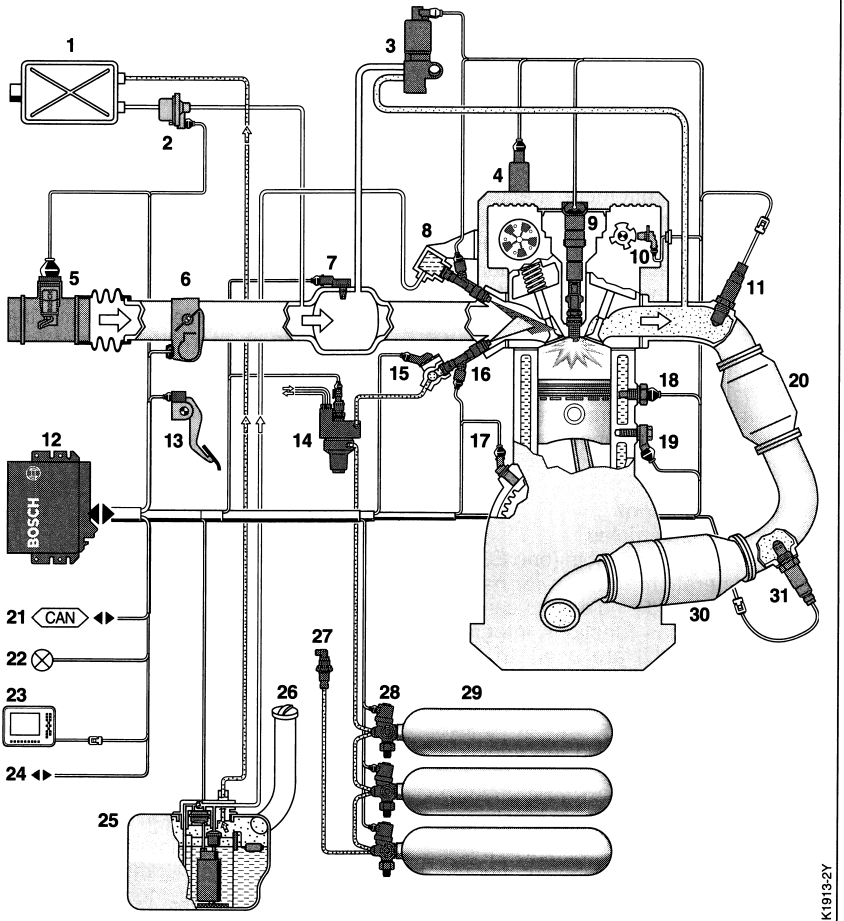
Natural-gas vehicles
Various vehicle manufacturers have in the meantime started to offer natural-gas versions of automobile models from the factory. This ensures that the larger CNG tanks are accommodated more efficiently and conveniently, thereby preventing – as is practically inevitable when these tanks are installed subsequently – the loss of luggage-compartment volume.

The latest information on natural-gas vehicles and the filling-station network in Germany and Europe can be found at the Internet addresses [1] and [2]. Natural-gas vehicles are, for the most part, bifuel vehicles which allow the driver to switch between natural-gas and gasoline operation. There are variants in the form of monovalent vehicles and the concepts referred to as "Monovalent plus", in which the engine is optimized for natural-gas operation so as to be able to utilize to optimum effect all the favorable properties of natural gas (higher knock resistance, lower CO_2 and pollutant emissions). "Monovalent plus" vehicles still feature a small gasoline tank (less than 15 *l*) so that they can continue to be driven in gasoline mode in the event of there being no natural-gas filling station nearby.

Storage of natural gas
Natural gas can be stored in compressed form at a pressure of 200 bar as CNG (compressed natural gas) or in liquid form at –162 °C as LNG (liquified natural gas). In view of the great expense involved in storing the gas in liquid form, storing the gas in compressed form at 200 bar has become the standard method. In spite of being stored at high-pressure, natural gas has a lower energy-storage density than gasoline such that a tank with four times the volume is required for the same energy content.

Figure 1: Spark-ignition engine with natural-gas or gasoline operation

1 Carbon canister, 2 Canister-purge valve, 3 Exhaust-gas recirculation valve,
4 Camshaft adjuster, 5 Hot-film air-mass meter,
6 Electronically actuated throttle valve (ETC),
7 Intake-manifold pressure sensor, 8 Gasoline rail with gasoline injectors,
9 Ignition coil with spark plug, 10 Camshaft sensor,
11 λ sensor before catalytic converter,
12 Bifuel Motronic ECU, 13 Accelerator-pedal module,
14 Natural-gas pressure-control module with integrated gas shutoff valve and high-pressure sensor,
15 Gas rail with natural-gas pressure and temperature sensor, 16 Natural-gas injector,
17 Crankshaft speed sensor, 18 Engine-temperature sensor, 19 Knock sensor,
20 Primary catalytic converter, 21 CAN interface,
22 Engine warning light, 23 Diagnosis interface, 24 Interface to immobilizer ECU,
25 Gasoline tank with integrated electric fuel pump, 26 Gasoline filler neck,
27 Natural-gas filler neck, 28 High-pressure shutoff valve on natural-gas tank,
29 Natural-gas tank, 30 Main catalytic converter,
31 λ sensor after primary catalytic converter.



Design

Components

Natural-gas automobiles are equipped almost exclusively with spark-ignition engines. They are extended to include the following natural-gas components (Figure 1, [3]):

- Natural-gas filler neck,
- Natural-gas tanks,
- High-pressure shutoff valves on the natural-gas tanks,
- Natural-gas pressure-control module with integrated gas shutoff valve and high-pressure sensor,
- Gas rail with natural-gas injectors and combined natural-gas low-pressure and temperature sensor.

Operating principle

The air drawn in by the engine is routed via the air-mass meter and the electronically actuated throttle valve to the intake manifold. From there, it enters the combustion chamber through the intake valves (Figure 1). The gas stored in the natural-gas tank at 200 bar flows through a high-pressure shutoff valve on the tank into the natural-gas pressure-control module, which lowers the gas pressure to a constant operating pressure of approx. 5 to 10 bar (absolute). It then flows through a flexible low-pressure line into a common gas rail that supplies the natural-gas injectors.

Engine management for natural-gas vehicles

Both double-ECU concepts (one ECU for gasoline operation and one for natural-gas operation) and single-ECU concepts (gasoline and gas functions integrated into a single ECU) are used. In some bifuel vehicles, the driver can use a switch to select between gas and gasoline operation. In most models, the switchover occurs automatically and the vehicle is driven in natural-gas mode until the tank contents are used up. Then the vehicle is switched over automatically to gasoline mode.

The high-temperature sensor mounted on the pressure-control module supplies the engine-management system with information on the current fill level of the natural-gas tank and also facilitates leakage diagnostics. The combined natural-gas low-pressure and temperature sensor fitted on the gas rail enables the engine-management system to correct the valve injection timing in such a way that a stoichiometric mixture can be pilot-controlled in the intake manifold even if the gas density fluctuates. The engine-management system uses an adaptation algorithm to adapt to changing gas qualities.

The other sensors and actuators of the engine-management system are mostly identical to those in a gasoline engine.

Mixture formation

Virtually exclusively bifuel and "Monovalent plus" vehicles are available today. These vehicles can operate both on natural gas and on gasoline. Either manifold injection or gasoline direct injection may be used to inject the gasoline.

Natural-gas manifold injection

In most natural-gas engines, the gas is injected, as with gasoline injection, into the intake manifold. The low-pressure gas rail supplies the injectors, which inject natural gas at intermittent intervals into the intake manifold. Mixture formation is improved in comparison with gasoline injection, as natural gas, because of the completely gaseous supply of fuel, does not condense on the intake-manifold walls and does not deposit a film there. This has a positive effect on emissions, particularly during the warm-up phase.

Non-supercharged engines that are fueled by natural gas generally have a 10 to 15% lower output. This is due to the lower volumetric efficiency attributed to displacement of the inducted air caused by the injected natural gas. The vehicle engines can however be optimized specially for natural gas as the fuel. The extremely high knock resistance of natural gas (up to 130 RON) provides for higher engine compression and is ideally suited to supercharging. In this way, the volumetric-efficiency losses can even be overcompensated. Combined with reduced piston displacement (downsizing), efficiency increases due to additional dethrottling and reduced friction.

Natural-gas direct injection
Natural gas can also be injected directly into the combustion chamber. In this way, the volumetric-efficiency losses can be completely avoided and in turbocharged engines the supercharging efficiency, especially at low speeds, can be improved by scavenging. Scavenging involves engine operation with extended overlapping of the intake- and exhaust-valve opening times. The drawbacks of natural-gas direct injection are the greater complexity of the injectors (tightness, temperature stability) and the fact that, due to the necessary higher natural-gas injection pressure dictated by the concept, the tank cannot be drained as far as in engines with natural-gas manifold injection.

Exhaust emissions
Compared with gasoline engines, natural-gas vehicles are characterized by approximately 25 % lower CO_2 emissions. The reason for this is the better hydrogen/carbon ratio (H/C ratio) of almost 4:1 (gasoline: approximately 2:1). This results in the creation of more water and less CO_2 during combustion.

Apart from the virtually particulate-free combustion, in conjunction with a three-way closed-loop catalytic converter, only very low levels of the pollutants NO_x, CO and HC are produced. As a rule, the catalytic converter has a higher precious-metal load than is the case in gasoline engines, enabling it to better convert the HC emissions in the exhaust gas that mainly consist of chemically stable methane and to counteract the higher "light-off temperature" (minimum temperature of the catalytic converter for conversion) for natural gas. Methane is, however, classified as nontoxic.

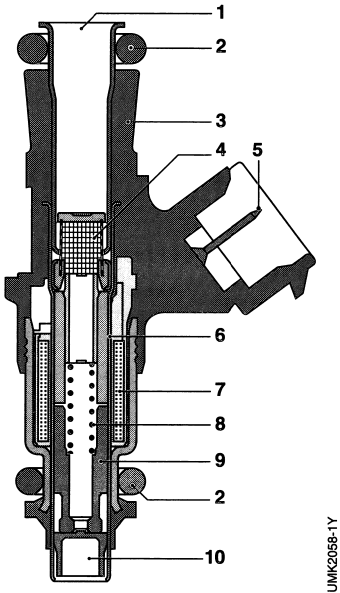
Natural-gas vehicles comply with the current exhaust-emission limits, while trucks and natural-gas buses also comply with the EEV limits (Enhanced Environmentally friendly Vehicle, [4]). The natural-gas engine offers distinct advantages over gasoline and diesel engines, also in particular with regard to the non-limited pollutant emissions.

Components

Natural-gas injector for manifold injection
Function
In order to supply the internal-combustion engine with gaseous fuel, it is necessary to meter through the natural-gas injectors a much greater volume of gas than gasoline in a conventional gasoline engine. This requirement places specific demands on the design of the natural-gas injector, which must be adapted to the greater gas volume in its cross-sections. Even the high flow velocities that are encountered call for a special form of flow routing in order to reduce pressure losses in the injector.

In highly supercharged engines, the intake-manifold pressure can rise up to 2.5 bar (absolute). In order to suppress the influence of the intake-manifold pres-

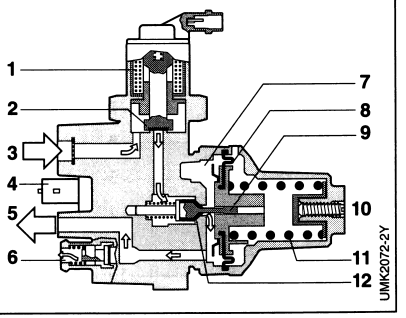
Figure 2: Natural-gas injector (Bosch)
1 Pneumatic port, 2 Sealing ring,
3 Valve housing, 4 Filter strainer,
5 Electrical connection, 6 Sleeve,
7 Solenoid coil, 8 Valve spring,
9 Solenoid armature with elastomer seal,
10 Valve seat.



sure on the mass flow, it is necessary for the pressure ahead of the nozzle at the narrowest cross-section (throttling point) accepted as the nozzle above critical to be at least twice as high as the maximum intake-manifold pressure (pressure after the nozzle). The gas then flows at the speed of sound, irrespective of the absolute pressure after the nozzle. The mass flow is therefore not influenced by the variable intake-manifold pressure. While possible pressure losses ahead of the throttling point are taken into account, a minimum operating pressure of 7 bar (absolute) has proven to be advantageous.

Design and operating principle
The solenoid armature (Figure 2) is guided in a sleeve. The armature has fuel flowing through it on the inside and has an elastomer seal at the discharge end. This seal closes on the valve seat and thereby seals off the fuel supply from the intake manifold. When energized, the solenoid coil effects the necessary force to lift the solenoid armature and open the metering cross-section (throttling point in the valve seat). When the coil is at zero current, the valve spring holds the injector closed.

Figure 3: Pressure-control module
1 Solenoid coil of electromagnetic gas shutoff valve,
2 Seal,
3 High-pressure inlet with sinter filter,
4 High-pressure sensor,
5 Low-pressure outlet,
6 Safety pressure-relief valve,
7 Low-pressure chamber,
8 Pressure diaphragm with mounting,
9 Control rod,
10 Adjusting screw,
11 Spring,
12 Control plunger and throttling orifice.



Flow-optimized geometry
Thanks to design measures in the routing of the flow, the pressure loss is minimized ahead of the throttling point in order to ensure the greatest possible mass flow. Furthermore, the narrowest cross-section and thus the throttling point is located at the discharge end after the seal. The speed of sound prevails here such that the valve represents approximately an ideal nozzle in physical terms.

Sealing geometry
The natural-gas injector is fitted with an elastomer seal and is similar in terms of its seal-seat geometry to shutoff valves for pneumatic applications. The elastomer increases the tightness of the seal against metallic needle valves.

Damping in the elastomer also prevents "rebounding", i.e. a repeated, unwanted opening of the solenoid armature during the closing operation, and thus increases metering precision.

Pressure-control module
Function
The function of the pressure-control module is to reduce the pressure of the natural gas from tank pressure (up to 200 bar) to the nominal operating pressure (typically 5 to 10 bar). At the same time, the operating pressure should be kept constant within specific tolerances through all operating states or, depending on the engine's operating point, regulated to a specified pressure. The typical operating pressure of present-day system is usually about 5 to 10 bar (absolute). There are also systems which operate at pressures starting from 2 bar ranging up to 11 bar.

Design
Mechanical version:
Today, mainly mechanical diaphragm- or plunger-type pressure regulators are used. Pressure reduction is effected by means of throttle action and can occur either in one single stage or in several stages.

Figure 3 shows the sectional view of a single-stage diaphragm-type pressure regulator. A sinter filter (pore size approx. 40 µm), a gas shutoff valve, and a high-pressure sensor are integrated into the pressure-control module on the high-pressure side. The sinter filter is designed to retain solid particulates in the gas flow. The gas shutoff valve enables the gas flow to be interrupted, for example when the vehicle is stopped. The high-pressure sensor is used to determine the natural-gas supply in the tank and for diagnostics.

A pressure-relief valve is mounted on the pressure regulator on the low-pressure side. In the event of a fault in the pressure regulator, this valve prevents damage to components in the low-pressure system.

In normal control operation, the gas cools down very dramatically as it expands as a result of the Joule-Thompson effect. In order to prevent freezing, the pressure-control module has an integrated heater, which is connected to the vehicle's heating circuit (not shown in the figure).

The controlled operating pressure is dependent on the diaphragm size and the initial spring tension. To adjust the pressure level, the initial spring tension is set and sealed by means of an adjusting screw in the factory.

Electromechanical version:
As well as purely mechanical pressure-control modules there are also electromechanical versions. These typically consist of an initial mechanical pressure-control stage, which reduces the tank pressure to a mean pressure of approximately 20 bar. In a second stage the pressure is further reduced by an electromagnetically actuated control valve to the operated pressure specified by the engine-management system, and electronically regulated. The advantages of electromechanical pressure regulators are the increased pressure-control precision and the variability of the control pressure. In this way, the injection pressure can be decreased and thus the metering precision of the natural-gas injectors increased in engine low-load phases, e.g. at idle.

Operating principle of mechanical pressure-control module
The gas flows from the high-pressure side through a variable throttling orifice into the low-pressure chamber, where the diaphragm is situated. The diaphragm controls the opening cross-section of the throttle via a control rod. When the pressure in the low-pressure chamber is low, the diaphragm is forced by the spring in the direction of the throttling orifice, which opens to allow the pressure to increase on the low-pressure side. In the event of excessive pressure in the low-pressure chamber, the spring is compressed more sharply and the throttling orifice closes. The decreasing cross-section of the throttling orifice reduces the pressure on the low-pressure side. In stationary operation, the throttling cross-section required for the operating pressure is set, while the pressure in the low-pressure chamber remains extensively constant.

Minimal operating-pressure fluctuations in the event of load changes testify to the quality of the pressure regulator.

References
[1] www.erdgas.info/erdgas-mobil/.
[2] www.gibgas.de.
[3] T. Allgeier, J. Förster: 2nd International CTI Forum, Stuttgart, March 2007; Einspritzsysteme – Motormanagementsystem und Komponenten für Erdgasfahrzeuge.
[4] www.bmub.bund.de/themen/luft-laerm-verkehr/luftreinhaltung/eev-standard/.

Engines fueled by alcohol

Application

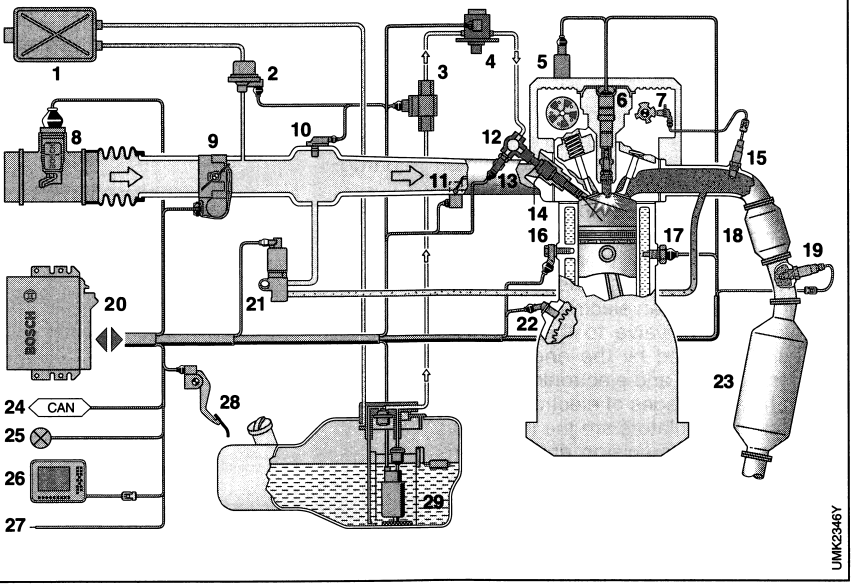
Motivation

Ever-shrinking fossil-fuel resources and the heightened demand for reduced carbon-dioxide emissions necessitate a rethink of tried-and-tested drive concepts. Rising crude-oil prices make alternative energies increasingly economical and there is the increasing public perception that sustained mobility also requires a sustained energy basis.

In Brazil, where ethanol has been used as a fuel since the 1980s and is available as E 24 and E 100 at filling stations (Table 1), approximately 50 % of fuel demand is covered. Alongside this, the USA and Europe have in the last few years increased their activities to promote E 85 as an alternative fuel type. Thus, aside from the potential to reduce CO_2, the focus is also aimed at securing reliability of supply and making fuel supply independent of fossil sources and imports.

Figure 1: Gasoline engine for flex-fuel operation

1 Carbon canister, 2 Tank vent valve, 3 Ethanol sensor (optional),
4 High-pressure pump, 5 Camshaft adjuster, 6 Ignition coil with connected spark plug,
7 Camshaft phase sensor, 8 Air-mass meter, 9 Throttle device (ETC),
10 Intake-manifold pressure sensor, 11 Swirl control valve,
12 Fuel rail, 13 High-pressure sensor, 14 High-pressure injector,
15 λ sensor, 16 Knock sensor, 17 Engine-temperature sensor, 18 Primary catalytic converter,
19 λ sensor, 20 Engine ECU, 21 Exhaust-gas recirculation valve,
22 Engine-speed sensor, 23 Main catalytic converter, 24 CAN interface,
25 Malfunction indicator lamp, 26 Diagnosis interface, 27 Interface to immobilizer,
28 Accelerator-pedal module, 29 Fuel-supply module with low-pressure pump.



System description

Flex-fuel systems are engine-management and component systems which can be operated with any mixture ratios of gasoline and ethanol between two limit values. These mixture ratios are created in the fuel tank by pumping in fuels with varying ethanol content. On the one hand, the capacity of a system to detect varying ethanol content in the fuel and to adapt its behavior (e.g., injection and ignition) accordingly plays a role. On the other hand, the question of component suitability is very important in making safe, robust operation possible over the entire service life demanded and expected by vehicle users.

Figure 1 shows the system diagram of a gasoline engine for flex-fuel operation. It is to a large extent the same as a conventional gasoline engine.

Ethanol as an alternative fuel

Sustainability

The potential to reduce CO_2 emissions is established in part by the more favorable ratio of carbon to hydrogen content of alcohol fuels, but primarily by the positive CO_2 balance in the course of extraction from renewable raw materials. It is however important here to consider the source of the produced alcohol fuel. Thus, for example, the extraction of methanol from coal (by means of the CTL process, Coal to Liquid) has an extremely poor CO_2 balance, whereas the extraction of ethanol from parts of plants has a positive balance.

Markets for flex-fuel vehicles

Today, the admixture of methanol in appreciable quantities plays, on account of local factors, a role only in China. Moreover, methanol, because it is toxic and harmful to health, must be viewed as highly critical. In the following, flex-fuel operation refers to operation with ethanol fuels.

Table 1: Composition of ethanol fuels

Fuel	E0...E5	E 10	E 24	E 85	E 100
Maximum ethanol content	5 %	10 %	24 %	85 %	93 %
Maximum water content	<1 %	<1 %	1 %	1 %	7 %
Minimum gasoline content	95 %	90 %	76 %	15 %	0 %
Relative energy content referred to gasoline	100 %	97 %	91 %	70 %	61 %
Regional distribution	Europe	USA, EU being introduced	Brazil,	USA, Sweden, partly EU	Brazil

Table 2: Properties of ethanol

Characteristic	Unit	Gasoline	Ethanol	Effects
Energy density (per volume)	kJ/l	32,500	21,200	– greater injected fuel quantity required
Energy density (per mass)	kJ/kg	43,900	26,800	– critical for cold start
Air/fuel ratio	–	14.8	9.0	– more wall film + more power, better efficiency (greater air cooling)
Boiling point	°C	25 to 215	78	
Enthalpy of evaporation	kJ/kg	380 to 500	904	
Octane number	RON	>91	108	+ higher knock limit + more power, higher efficiency
H/C ratio	–	2.3	3	

Flex-fuel refers to mixtures of gasoline and ethanol in any variable ratio. Today, a distinction is drawn primarily between two mixture ranges of ethanol mixtures in the flex-fuel sector: Mixtures with mixture ratios between pure gasoline and E 85 (above all Europe – particularly in Sweden – and in North America) and between E 24 and E 100 in Brazil. The lower limit towards E 18 is currently being extended in Brazil.

Another specific feature is that E 100 fuel according to the Brazilian standard (Resolution ANP #19 – Technical Regulation, [1]) may contain up to 7.5 % water (EHC, Hydrated Ethanol Fuel), but in real terms in the field contains 10 to 15 % water. Also associated with this increased water content is increased conductivity and the increased presence of salts. This is the reason why there are two separate variants of a component for the E 85 and E 100 application cases and also specific adaptations are made in the electronic engine management. It is important here than an E 85 system is not suitable for E 100.

Essentially, flex-fuel systems are implemented with manifold and direct injection. In the case of direct injection it is possible in particular to use the high knock resistance of ethanol to good effect, which is advantageous with regard to the full-load quantity demand due to less fuel enrichment.

Properties of ethanol and resulting requirements

The specific properties of ethanol that deviate from regular gasoline are set out in Table 2. These impact on system performance and on the components in the system that come directly or indirectly into contact with fuel.

Important factors are the lower energy density (greater quantity demand per unit of air) and the deviating stoichiometry of ethanol caused by the chemical composition (oxygen is already contained in part), which results in dramatically increased fuel-system delivery and injected fuel quantities. Critical to the starting response at very low temperatures are the high boiling temperature of 78 °C and the high heat of evaporation compared with the low-boiling fractions of gasoline. However, for operation at high engine temperatures and at the knock limit, advantageous effects are also to be expected due to the high heat of evaporation and better knock resistance.

These mentioned properties of ethanol and their effects make it necessary to adapt the internal-combustion engine itself, but also the engine-management system and its components accordingly. It is also important that the ethanol fuels not be present as pure fuels with a fixed (and thus known) ethanol content in the fuel tank, but that the mixing of different ethanol fuels lead to any mixture ratios. The mixtures can moreover also be supplied by regionally and seasonally deviating ethanol admixtures by the filling stations and in the USA by "blender pumps" in the system. Thus, a reliable and accurate "on-board recognition" of the current ethanol content is required.

The higher corrosivity of ethanol must be taken into account for the components installed in a system. To achieve evaporation limit values the higher permeation tendency of ethanol molecules through pipe and component materials plays a significant role.

Flex-fuel concepts for the markets

Vehicle manufacturers pursue different engine concepts and marketing strategies for the regions. With regard to the degree of adaptation of the systems there are therefore different aspects which conform to the aimed-for strategy.

Markets

USA market

In the US market as a matter of priority only those adaptations which are necessary to be able to offer the vehicles as flex-fuel vehicles are carried out. This involves the fundamental compatibility for flex-fuel without specifically exploiting the advantageous ethanol properties. In real vehicle operation it is used only occasionally with higher and variable ethanol contents.

European market

In contrast, a different concept is pursued in Europe whereby both the increased efficiency of the engine and the enhanced performance in E85 operation are exploited as advantages, and are emphasized in the marketing of the vehicle.

Brazilian market

Only in Brazil is it possible up to now to observe the last logically consistent step where the engine manufacturers also increase the compression ratios of the engines since in this market already gasoline generally contains at least 24 % ethanol and consequently a higher octane number (and thus a higher knock resistance) is obtained.

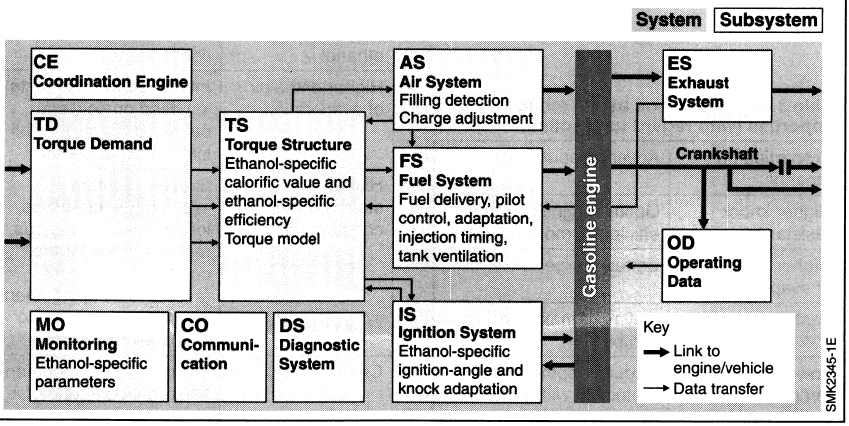
System adaptations

Extensive adaptations are required in the management system of the engine ECU and its functions. Specifically, the following engine-management functions are affected (Figure 2):

- Air system: e.g. filling detection and cylinder-charge control.
- Fuel system: e.g. quantity pilot control and injection timing, mixture adaptation.
- Exhaust system: e.g. enabling of mixture control.
- Ignition system: e.g. moments of ignition, ignition energy.
- Torque patterns: e.g. deviating efficiency with regard to λ and ignition.
- Diagnostic system and monitoring; e.g. fuel-quality detection, diagnostic thresholds.

When it comes to system adaptation the deviating properties of ethanol with regard to gasoline must be taken into account in order to avoid diminished system performance or to achieve improved perfor-

Figure 2: Software structure of engine management for flex-fuel operation



mance in partial ranges. The properties listed in Table 3 can be used to good effect, while those listed in Table 4 necessarily require an adaptation.

In part, however, new developments of functionalities are required. The following are mentioned here by way of example:
- Special cold-start strategies for achieving emission limits at lower temperatures (e.g. –7 °C) and for fundamentally ensuring engine starting at very low temperatures (–30 °C).
- Detection of a more pronounced impingement of fuel in the engine oil and interception of increase degassing during warm-up.
- Detection of alternating ethanol contents in the fuel tank and diagnosis of ethanol detection (sensor-based or sensorless).

Flex-fuel components

Flex-fuel systems are virtually identical in design to gasoline injection systems; however, the demands imposed on the components differ. With regard to component suitability, it is necessary to consider both the fuel-carrying components and those components that potentially come into contact with fuel vapors or combustion gases of ethanol. An overview of the components to be adapted in a flex-fuel system is provided in Table 5.

Changes in the components must be made with regard to the following influences:
- corrosion of metallic materials,
- swelling or embrittlement of plastic and rubber parts,
- functional capability of gaskets,
- increased permeation of fuel molecules through fuel-carrying components.

This is done specifically by means of the following measures:
- use of special alloys (e.g. stainless steels),
- surface coating of metallic materials, particularly in mechanically stressed contact pairs (above all on AL materials, e.g. Nikasil coatings: nickel and silicon carbide),
- use of adapted plastic and rubber compounds,

Table 3: Advantageous use of ethanol properties (with regard to gasoline)

Properties of ethanol	Advantageous use
Higher knock resistance	Optimum ignition, better efficiency, more power
Higher enthalpy of evaporation	Better charge
Faster combustion	Optimum maximum combustion
Lower combustion temperatures	Reduced component protection with lower consumption

Table 4: Disadvantageous consequences of ethanol properties (with regard to gasoline)

Properties of ethanol	Disadvantageous consequences
Higher enthalpy of evaporation	Increased requirements imposed on cold start below 10 °C, larger wall film in cold operation
Higher stoichiometric consumption	Greater required injected fuel quantities, longer injection periods, repercussions for the combustion process, greater fuel impingement in the engine oil during cold start
Changed H/C ratio	More water impingement, later dew-point end of λ sensor

Table 5: Overview of ethanol-specific components

Component	Adaptation
Fuel-supply module	Fuel pump for E 85 or E 100, insulated electrical connections, Fuel-level sensor ethanol-resistant
Pressure regulator	Ethanol-resistant version
Fuel rail	Ethanol-resistant version
Intake-manifold fuel injector	Extended metering range
High-pressure injector	Stainless-steel versions, adapted metering range
High-pressure pump	Stainless-steel versions, adapted metering range

- selection of specific gasket and filter materials,
- design adaptations,
- electric insulation of electrical contacts in the fuel tank (electric fuel pump, fuel-level sensor) for E 100 due to the water content of up to 15 %.

Increased requirements with regard to greater flow rates must be taken into account for specific components. This is particularly important in relation to the fuel injectors since the maximum flow rate is necessary for the increased demand in ethanol operation, but still the minimum metering quantity occurs in gasoline operation. This adaptation is effected by the selection of a greater flow rate from the range of variants. The unchanged minimum quantity is implemented by a shortened activation duration.

References

[1] http://www.itecref.com/pdf/Brazilian_ANP_Fuel_Ethanol.pdf.
[2] A. Böge: Handbuch Maschinenbau, 22nd Ed., Verlag Springer Vieweg, 2014.
[3] K.-H. Grote and J. Feldhusen (Editors.): Dubbel: Taschenbuch für den Maschinenbau, 24th Ed., Verlag Springer Vieweg, 2014.
[4] GVR Gas Vehicle Report, June 2012.

Management for diesel engines

Description of the engine-management system

In the diesel combustion process, the fuel is injected directly into the combustion chamber at a nozzle pressure ranging from 200 to more than 2,200 bar. Depending on the combustion process, fuel is injected into a prechamber in indirect-injection engines (at a relatively low pressure of less than 350 bar). State-of-the-art technology is the direct-injection process, in which fuel is injected into the non-subdivided combustion chamber at high pressure up to over 2,500 bar.

Diesel-engine management

Requirements

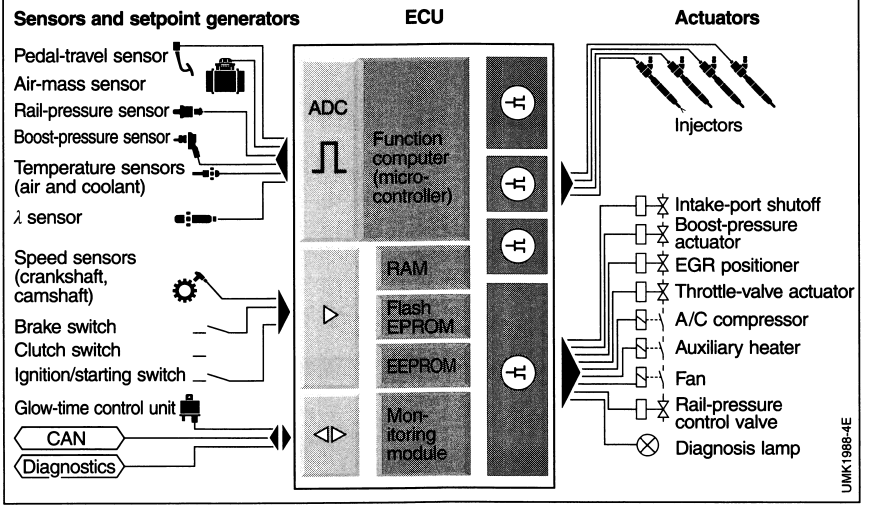
The power output P from a diesel engine is determined by the available clutch torque and the engine speed. The clutch torque is produced from the torque generated by the combustion process, reduced by the friction torque and the charge-cycle losses, and the torque required for operating the auxiliary systems driven directly by the engine. The combustion torque is generated in the power cycle and is determined by the following variables if the excess air is sufficient: the supplied fuel mass, the start of combustion determined by the start of injection, and the injection and combustion processes.

In addition, the maximum speed-dependent torque is limited by the smoke emissions, the cylinder pressure, the temperature load of different components, and the mechanical load of the complete drivetrain.

Primary function of engine management

The primary function of engine management is to adjust the torque generated by the engine or, in some applications, to adjust a specific engine speed within the permitted operating range (e.g. idling).

Figure 1: Components for Electronic Diesel Control (EDC) on the example of a common-rail system



In a diesel engine, exhaust-gas treatment and noise suppression are performed to a great extent inside the engine, i.e. by controlling the combustion process. This, in turn, is performed by engine management by changing the following variables:
- cylinder charge,
- temperature control of the cylinder charge via intercooling (charge-air cooling) and EGR cooling,
- composition of the cylinder charge (exhaust-gas recirculation),
- charge motion (intake swirl),
- start of injection,
- injection pressure, and
- rate-of-discharge curve control (e.g. pilot injection, divided fuel injection).

Until the 1980s fuel injection in vehicle engines, i.e. injected fuel quantity and start of ignition, was governed exclusively by mechanical means. In an in-line fuel-injection pump, for example, the injected fuel quantity is varied as a function of load and speed by rotating the pump plunger, which has an angled helix. In the case of mechanical control, the start of injection and start of delivery are adjusted by flyweight governors (speed-dependent), or hydraulically by pressure control as a function of speed and load.

The emission-control legislation demands very precise regulation of injected fuel quantity and start of injection, depending on variables such as engine temperature, engine speed, load, or altitude. This can only be provided effectively by electronic control. Now electronic control has fully replaced mechanical control. This is the only form of control that permits continuous monitoring of the emission-related function of the fuel-injection system. Legislation also requires on-board diagnosis in some applications.

Control of injected fuel quantity and start of injection is performed in EDC systems (Electronic Diesel Control) by means of low- or high-pressure solenoid valves or other electrical actuators. Rate-of-discharge curve control within an injection sequence, i.e., dividing the fuel quantity into several partial injections, is performed indirectly via a servo valve in the solenoid-valve or piezo inline injector.

Electronic Diesel Control

Electronic control of a diesel engine (Figure 1, Electronic Diesel Control, EDC) enables precise and differentiated modulation of fuel-injection parameters. This is the only means by which a modern diesel engine is able to satisfy the many demands placed upon it.

System overview

Requirements

The lowering of fuel consumption and pollutant emissions (NO_x, CO, HC, particulates) combined with simultaneous improvement of engine power output and torque are the guiding principles of current development work on diesel-engine design. In recent years this has led to an increase in the popularity of the direct-injection (DI) diesel engine, which uses much higher fuel-injection pressures than indirect-injection (IDI) engines with whirl or prechamber systems. In addition, diesel-engine development has been influenced by the high levels of comfort and convenience demanded in modern cars. Noise levels, too, are subject to more and more stringent requirements. As a result, the performance demanded of fuel-injection and engine-management systems has also increased, specifically with regard to
- high injection pressures,
- rate shaping,
- pre-injection and, if necessary, post-injection
- variation of injected fuel quantity, charge-air pressure, and start of injection to suit operating conditions,
- temperature-dependent excess-fuel quantity for starting,
- load-independent idle-speed control,
- controlled exhaust-gas recirculation,
- cruise control, and
- tight tolerances for start of injection and injected fuel quantity and maintenance of high precision over the service life of the system (long-term performance).

Conventional mechanical governing of engine speed uses a number of adjusting mechanisms to adapt to different engine operating conditions. Nevertheless,

it is restricted to a simple engine-based control loop and there are a number of important influencing variables that it cannot take account of or cannot respond quickly enough to. As demands have increased, what was originally a straightforward system using electric actuators has developed into present-day Electronic Diesel Control (e.g. distributor-type injection pump), a complex electronic engine-control system capable of processing large amounts of data in real time. And as a result of the increasing integration of electronic components, the control-system circuitry can be accommodated in a very small space.

Operating principle

Electronic Diesel Control is capable of meeting the requirements listed above as a result of microcontroller performance that has risen considerably in the last few years.

In contrast to diesel-engine vehicles with conventional in-line or distributor injection pumps, the driver of an EDC-controlled vehicle has no direct influence, for instance through the accelerator pedal and Bowden cable, upon the injected fuel quantity. Instead, the injected fuel quantity is determined by a number of influencing variables. These are, for example, driver command (accelerator-pedal position), operating status, engine temperature, intervention by other systems (e.g. traction control), and effects on exhaust emissions.

Start of injection can also be varied. This demands a comprehensive monitoring concept that detects inconsistencies and initiates appropriate actions in accordance with the effects (e.g. torque limitation or limp-home mode in the idle-speed range). Electronic Diesel Control therefore incorporates a number of control loops.

Electronic Diesel Control allows data communication with other electronic systems, such as the Traction Control System (TCS), Electronic Transmission Control (ETC), or driving-dynamics control. As a result, the engine-management system can be integrated in the vehicle's overall control system network, thereby enabling functions such as reduction of engine torque when the automatic transmission changes gear, regulation of engine torque to compensate for wheel spin.

The EDC system is fully integrated in the vehicle's diagnosis system. It meets all OBD (On-Board Diagnostics) and EOBD (European OBD) requirements.

System blocks

Electronic Diesel Control is divided into three system blocks (Figure 1). Sensors and setpoint generators detect operating conditions (e.g. engine speed) and setpoint values (e.g. switch position). They convert physical variables into electrical signals.

The engine ECU processes the information from the sensors and setpoint generators in its function computer using open- and closed-loop control algorithms. It controls the actuators by means of electrical output signals. In addition, the ECU acts as an interface to other systems and to the vehicle diagnosis system.

Actuators convert electrical output signals from the control unit into mechanical parameters (e.g. the solenoid valve for the fuel-injection system).

Data processing

The main function of Electronic Diesel Control is to control the injected fuel quantity and the injection timing. The common-rail fuel-injection system also controls injection pressure.

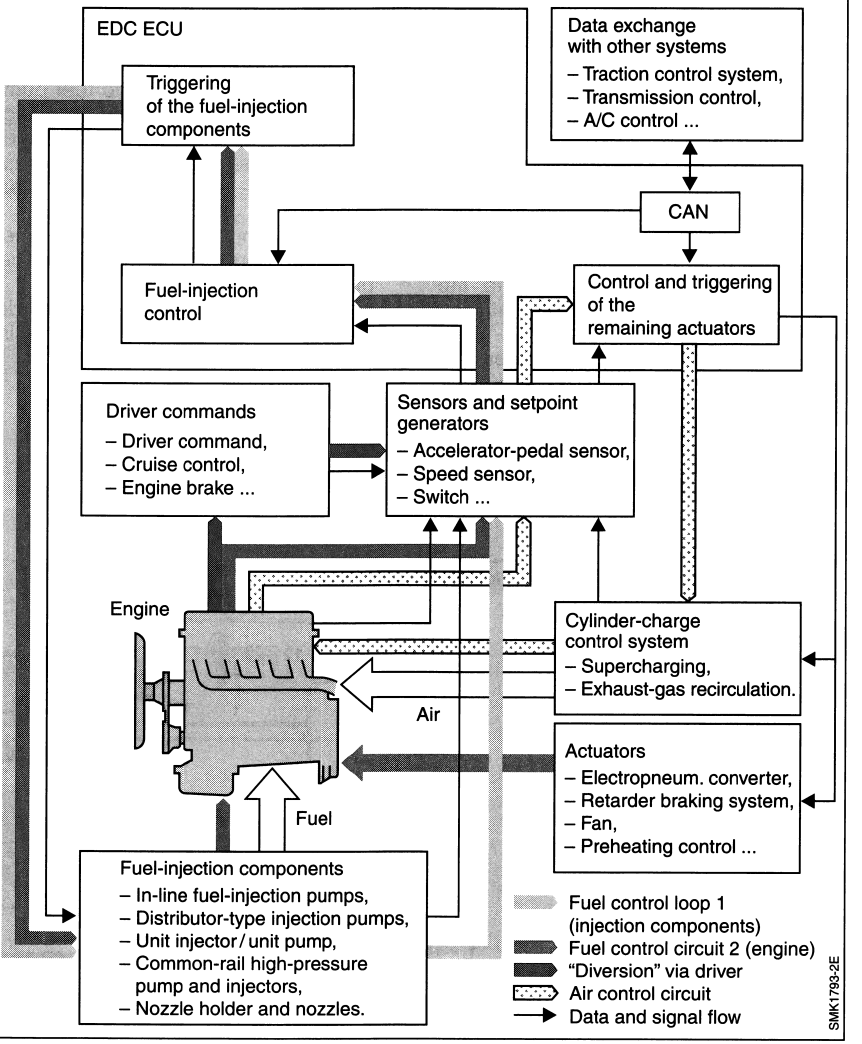
The ECU analyzes the incoming signals from the sensors and limits them to the permissible voltage level. Some input signals are also checked for plausibility. Using these input data together with stored program maps, the function computer calculates injection timing and its duration. This information is then converted to a signal characteristic which is aligned to the engine's piston strokes. This calculation program is termed the "ECU software". For all components to operate efficiently, the EDC functions must be precisely matched to every vehicle and every engine. This is the only way to optimize component interaction (Figure 2).

The required degree of accuracy together with the diesel engine's outstanding dynamic response requires high-level computing power. The output signals are

applied to output stages, which provide adequate power for the actuators (e.g., injectors, EGR positioner, boost-pressure actuator). Apart from this, a number of other auxiliary-function components (e.g. glow relay and air-conditioning system) are triggered.

Faulty signal characteristics are detected by output-stage diagnostic functions for the solenoid valves. Furthermore, signals are exchanged with other systems in the vehicle via the interfaces. The engine ECU monitors the complete fuel-injection system as part of a safety strategy.

Figure 2: Basic sequence of Electronic Diesel Control



Low-pressure fuel supply

The function of the fuel-supply system is to store and filter the required fuel and to make fuel available to the fuel-injection installation at a specific supply pressure under all operating conditions. The fuel return flow is cooled in some applications. The fuel supply system comprises the following essential components (Figures 1 to 3): the fuel tank, the preliminary filter (not for unit-injector system for passenger cars), the presupply pump (optional, also in-tank pump for passenger cars), the fuel filter, the fuel pump (low pressure), the pressure-control valve (overflow valve), the fuel cooler (optional), and low-pressure fuel lines.

Individual components can be combined to form assemblies (e.g. fuel-supply pump with pressure limiter). In axial- and radial-piston distributor pumps and partially in the common-rail system, the fuel-supply pump is integrated into the high-pressure pump.

Essentially, the fuel-supply system differs greatly, depending on the fuel-injection system used, as the following figures for common rail (Figure 1), unit injector (Figure 2) and radial-piston pump (Figure 3) show.

Fuel supply and delivery

The function of the fuel pump in the low-pressure stage is to supply the high-pressure components with sufficient fuel in every operating state, with a low level of noise, at the required pressure, and over the entire service life of the vehicle. Different types are used, depending on the field of application.

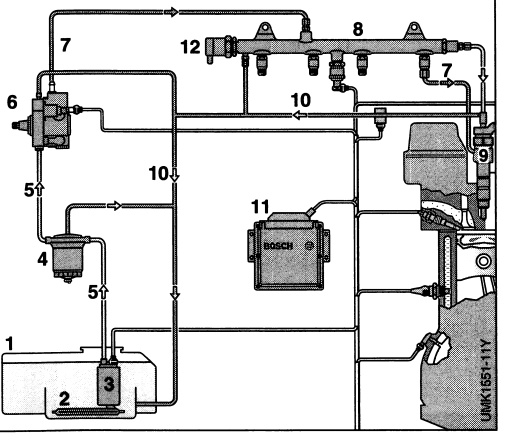
Electric fuel pump

The structure of the electric fuel pump (Figure 4) is comparable to the design variations of the pumps used on spark-ignition engines (see Electric fuel pump). However, there are some different features:

- Roller-cell pumps are usually used for diesel applications (see Roller-cell pump),
- The electric motor of the pump is designed with carbon commutators instead of copper commutators.

Figure 1: Fuel system of a fuel-injection installation with common rail for passenger cars

1 Fuel tank,
2 Preliminary filter,
3 Fuel pump,
4 Fuel filter,
5 Low-pressure fuel lines,
6 High-pressure pump,
7 High-pressure fuel lines,
8 Rail,
9 Injector,
10 Fuel return line,
11 ECU,
12 Pressure-control valve.



Use of electric fuel pumps:
- Optional for distributor injection pumps (only in event of long fuel lines or great difference in height between fuel tank and fuel-injection pump)
- Unit-injector system (passenger cars)
- Common-rail system (passenger cars)

Gear pump

The gear pump is mounted directly onto the engine or, in the case of a common rail, it is integrated into the high-pressure pump. It is driven mechanically by way of a clutch, a gear wheel, or a toothed belt.

The primary components are two intermeshing, counter-rotating gear wheels (Figure 5) which deliver the fuel into the

Figure 2: Fuel system of a fuel-injection installation with unit-injector system for passenger cars

1 Fuel tank with presupply pump and prefilter,
2 Fuel cooler,
3 Electronic control unit,
4 Fuel filter,
5 Fuel feed line,
6 Fuel return line,
7 Tandem pump as fuel pump,
8 Fuel-temperature sensor,
9 Sheathed-element glow plug,
10 Injector.

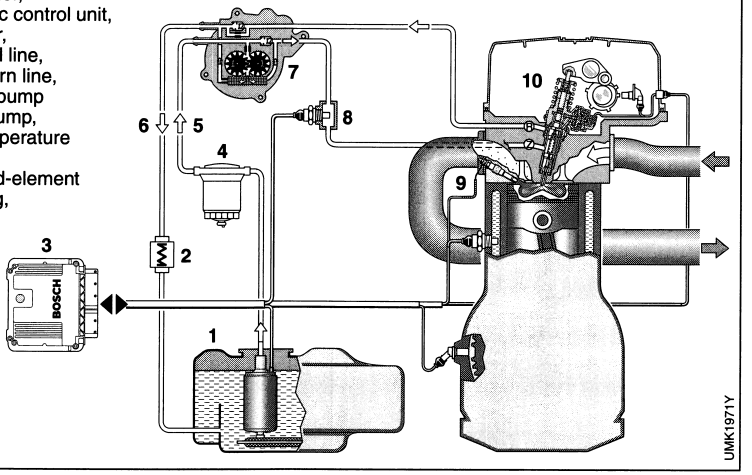
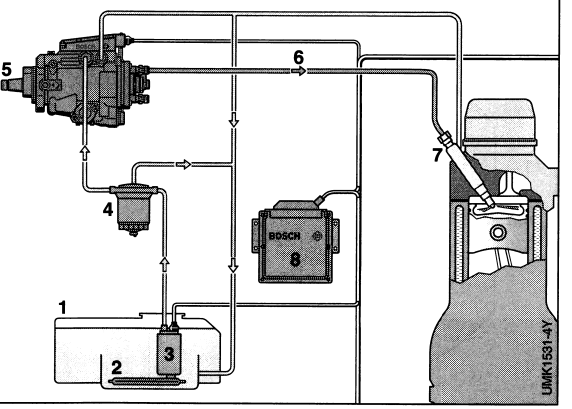


Figure 3: Fuel system of a fuel-injection installation with radial-piston fuel-injection pump for passenger cars

1 Fuel tank,
2 Preliminary filter,
3 Presupply pump,
4 Fuel filter,
5 Radial-piston distributor injection pump with integrated fuel pump,
6 High-pressure line,
7 Nozzle-holder assembly,
8 Control unit.



tooth spaces from the suction side to the pressure side. The line of contact of the gear wheels provides the seal between the suction and pressure sides and prevents fuel from flowing back.

The delivery quantity is roughly proportional to the engine speed. At high rotational speeds, therefore, the quantity is limited by throttle action on the intake side or by an overflow valve on the pressure side, which discharges the feed volume to the inlet inside the pump. This prevents unnecessarily high flows in the vehicle's conduit system.

Applications:
- For single-cylinder pump systems for commercial vehicles (unit-injector and unit-pump system, individual injection pump (PF))
- for common-rail systems (commercial vehicles, in part for passenger cars).

Figure 4: Single-stage electric fuel pump
A Pump element,
B Electric motor,
C Connecting cover.
1 Pressure side,
2 Motor armature,
3 Pump element,
4 Pressure limiter,
5 Suction side,
6 Non-return valve.

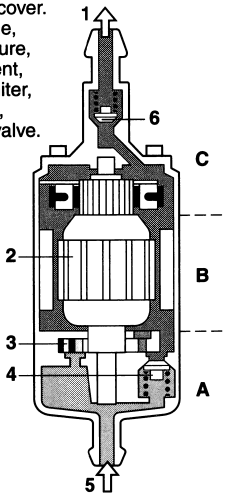
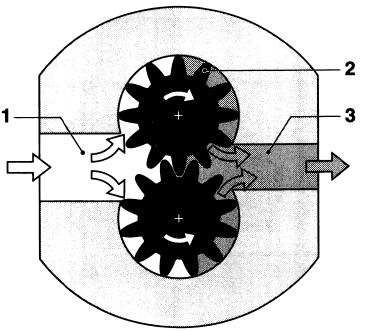


Figure 5: Gear fuel pump (diagram)
1 Suction side,
2 Drive gear,
3 Pressure side.

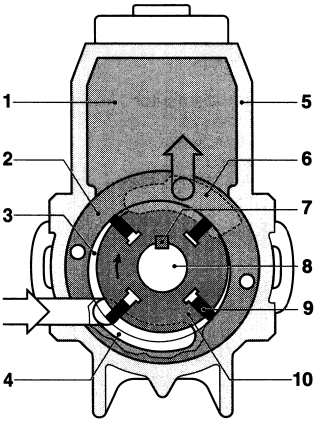


Vane-type supply pump

The vane-type supply pump (Figure 6) is mounted on the drive shaft in the distributor injection pump. The impeller is centered on the drive shaft and supported by a Woodruff key. An eccentric ring mounted in bearings in the housing surrounds the impeller.

The fuel passes through the inlet passage and a kidney-shaped recess into the space created by the impeller, the

Figure 6: Vane pump
1 Inner pump chamber,
2 Eccentric ring,
3 Sickle-shaped cell,
4 Fuel inlet (suction cell),
5 Pump housing,
6 Fuel outlet (pressure cell),
7 Woodruff key,
8 Drive shaft,
9 Vane, 10 Impeller.



vane and the eccentric ring. The centrifugal force generated by the motion of rotation presses the four vanes of the impeller outwards against the eccentric ring. The fuel between the underside of the vanes and the impeller supports this outward movement of the vanes. The rotary motion forces the fuel located between the vanes into the upper kidney-shaped recess and through a bore to the outlet.

Application: Integrated presupply pump in the distributor injection pump.

Locking-vane pump

In a locking-vane pump (Figure 7), springs press two locking vanes against a rotor. As the rotor turns, the volume is increased on the suction side and fuel is drawn into two chambers. The volume is reduced on the pressure side and fuel is delivered from the two chambers. The locking-vane pump delivers fuels at very low speeds.

Application: Unit-injector system for passenger cars.

Tandem fuel pump

The tandem fuel pump (Figure 8) is an assembly consisting of a fuel pump and a vacuum pump, for example, for the brake booster. From a functional perspective, these are two separate pumps, which, however, are driven by a common drive shaft. It is integrated into the engine cylinder head and driven by the engine camshaft.

The fuel pump itself is either a locking-vane pump or a gear pump, and even at low speeds (cranking speeds) delivers enough fuel to ensure that the engine starts reliably. The pump's delivery quantity is proportional to the rotational speed.

The pump contains a variety of valves and throttling orifices. The pump's maximum delivery quantity is limited by the suction throttling orifice so that not too much fuel is delivered. The pressure-relief valve limits the maximum pressure in the high-pressure stage. Vapor bubbles in the fuel feed are eliminated in the fuel-return throttling bore. If there is air in the fuel system (e.g. if the vehicle has been driven until the fuel tank is empty), the low-pressure control valve remains

Figure 7: Locking-vane pump (diagram)
1 Rotor,
2 Suction side (inlet),
3 Spring,
4 Locking vane,
5 Pressure side.

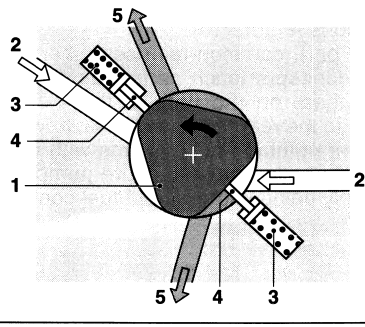


Figure 8: Fuel pump in a tandem pump
1 Return to tank,
2 Supply from tank,
3 Pump element (gear wheel),
4 Throttling bore,
5 Filter,
6 Suction throttle,
7 Pressure-relief valve,
8 Connection for pressure measurement,
9 Supply, injector,
10 Return, injector,
11 Non-return valve,
12 Bypass.

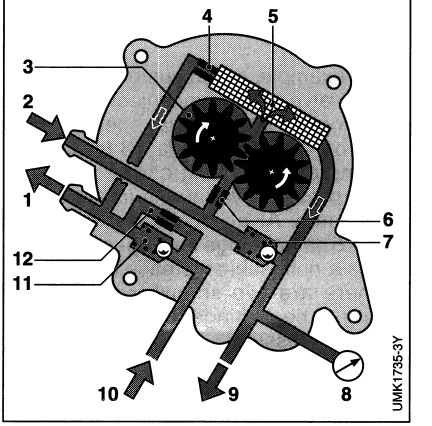


Figure 9: Pressure-control valve
1 Valve body,
2 Screw,
3 Compression spring,
4 Edge seal,
5 Accumulator plunger,
6 Accumulator volume,
7 Conical seat.

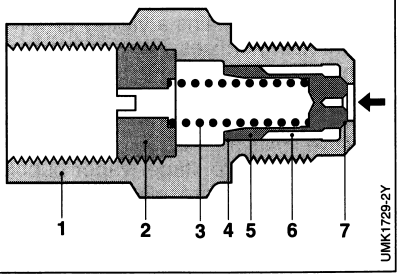
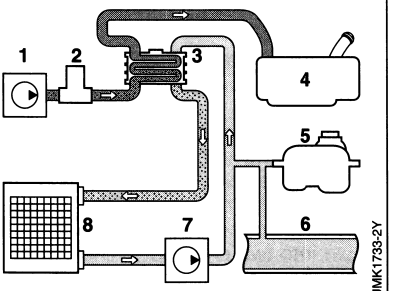


Figure 10: Fuel-cooling circuit
1 Fuel-supply pump,
2 Fuel-temperature sensor,
3 Fuel cooler,
4 Fuel tank,
5 Expansion tank,
6 Engine-cooling circuit,
7 Coolant pump,
8 Auxiliary cooler.



closed. The air is forced out of the fuel system through the bypass by the pressure of the pumped fuel.

Application: Unit-injector system for passenger cars.

Low-pressure control valve

The pressure-control valve (Figure 9, also called overflow restriction) is built into the fuel return, for common-rail usually into the high-pressure pump. It provides an adequate operating pressure in the low-pressure stage of the injection systems under all operating conditions and ensures that the pumps are filled uniformly.

For example, the accumulator plunger of the valve for the unit injector and unit pump opens at the initial opening pressure of 300 to 350 kPa (3 to 3.5 bar). A compression spring compensates for minor pressure variations in accumulator volume. At an opening pressure of 4 to 4.5 bar, an edge seal opens and produces a noticeable increase in flow rate.

There are two screws with different spring-force settings for presetting the opening pressure.

Fuel cooler

The high pressure in the injector of the unit-injector system for passenger cars and some common-rail systems heats up the fuel to such an extent that it has to be cooled before it flows back. The fuel returning from the injector flows through the fuel cooler (heat exchanger) and dissipates thermal energy to the coolant in the fuel-cooling circuit. This is separate from the engine-cooling circuit because the coolant temperature of an engine running at operating temperature is too high to cool the fuel. The fuel-cooling circuit is connected to the engine-cooling circuit at the compensating reservoir. This helps to keep the fuel-cooling circuit filled and compensates for any changes in volume caused by temperature fluctuations (Figure 10).

In part, common-rail uses other cooling concepts (such as fuel-air heat exchangers on the floor of the vehicle). Due to the reduced power loss, however, newer common-rail systems with quantity-controlled high-pressure pumps are largely being used without fuel coolers.

Diesel filter

Function

As in spark-ignition (SI) engines, it is essential in a diesel engine to ensure that the fuel system is protected against contamination. Contaminants are introduced into the system during refueling or can get into the tank via the tank ventilation and thus into the fuel itself. The function of the fuel filter is to reduce contamination from particulates so as to protect the fuel-injection system.

Compared with gasoline-injection systems, diesel fuel-injection systems require increased wear protection and finer filters to cater for the much higher injection pressures. Diesel fuel is also more heavily contaminated than gasoline.

Design

Diesel filters are designed as easy-change filters (Figure 1). Spin-on filters, in-line filters and nonmetal filter elements are widely used as replacement parts in aluminum, plastic or sheet-steel filter cases (to comply with increased requirements in crashes). Spiral vee-shape filter elements are the preferred choice. The diesel fuel filter is installed in the low-pressure circuit: in suction systems upstream of the fuel-supply pump and in pressure systems downstream of the electric fuel pump. The trend is now towards pressure systems.

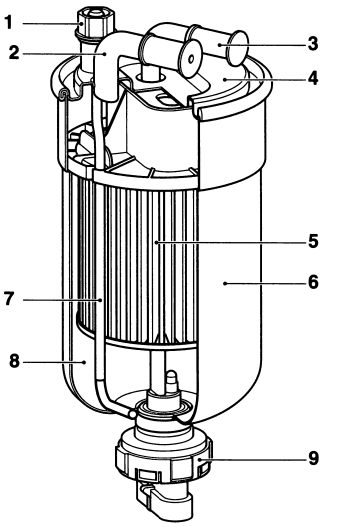
Requirements

In the last few years, the requirements for filter fineness have become more stringent owing to the introduction of common-rail systems with higher injection pressures and advanced unit-injector systems for passenger cars and commercial vehicles. Depending on the application (fuel contamination, periods of engine standstill), the new systems require a filtration efficiency of between 85 % and 98.6 % (particulate interval 3 to 5 µm, ISO/TS 13353 [1] and ISO 19438 [2]). The fuel filters fitted to the latest generation of automobiles must be capable of storing larger volumes of particulates due to the longer service intervals, and of separating superfine particulates efficiently. This can only be achieved by using special filter media, e.g. consisting of multiple layers made of synthetic microfiber. These filter media exploit the effects of a fine prefilter, and guarantee maximum particulate-retention capability by separating particulates inside each filter layer.

Typical replacement intervals today are between 60,000 and 90,000 km. These intervals are significantly shortened in markets where the diesel quality is poor, such as Eastern Europe, China,

Figure 1: Diesel line filter

1 Water drain plug,
2 Fuel inlet,
3 Fuel outlet,
4 Filter cover with galvanized coating,
5 Radial vee-form filter element with two-ply filter medium,
6 Compression-proof housing made from Galfan (hot-dip-modified sheet with a zinc-aluminum alloy coating on both sides),
7 Tube to drain,
8 Water accumulation chamber,
9 Water-level and temperature sensors.



India, and the USA. Likewise, when biodiesel is used, halving the replacement interval is recommended.

Water separation

A second essential function of the diesel filter is to separate emulsified and free water in order to prevent corrosion damage. An effective water separation of more than 93% at rated flow (testing as per ISO 4020 [3]) is particularly important for distributor injection pumps and common-rail systems. Water separation takes place by coalescence on the filter medium (formation of droplets due to the different surface tensions of water and fuel). Separated water collects in the chamber at the bottom of the filter housing (Figure 1). Conductivity sensors are used in some cases to monitor the water level. The water is drained manually by opening a water drain plug or pressing a pushbutton switch.

For extra heavy-duty requirements, it is beneficial to fit an additional preliminary filter with water separator on the suction or pressure side. Preliminary filters of this type are mostly used for commercial vehicles in countries where diesel quality is poor.

Modular additional functions

New-generation diesel fuel filters integrate modular additional functions such as, for example, fuel preheating to prevent paraffin clogging during winter operation. Fuel preheating can be effected electrically or by utilizing the hot fuel return flow from the engine. In the first instance, PTC heaters (Positive Temperature Coefficient) are installed in the filter. The second instance requires the fitting of a bimetal valve or a wax expansion element which opens at low temperatures and allows fuel to flow back to the filter.

Further additional functions are maintenance indication by way of differential-pressure measurement and filling and venting devices.

References

[1] ISO/TS 13353:2002: Diesel fuel and petrol filters for internal combustion engines – Initial efficiency by particle counting.

[2] ISO 19438:2003: Diesel fuel and petrol filters for internal combustion engines – Filtration efficiency using particle counting and contaminant retention capacity.

[3] ISO 4020:2001: Road vehicles – Fuel filters for diesel engines – Test methods.

Common-rail injection system

System overview

Requirements

The demands placed on diesel-engine fuel-injection systems are continually increasing. Higher injection pressures, faster injector switching times, and a variable rate-of-discharge curve modified to the engine operating state have made the diesel engine economical, clean, and powerful.

The main advantage of the common-rail system (CR) over conventional fuel-injection systems is its ability to vary injection pressure and timing over a broad scale. This was achieved by separating pressure generation (in the high-pressure pump) from the fuel-injection system (injectors). The rail here acts as a pressure accumulator (Figure 1).

The common-rail system is a highly flexible system for adapting fuel injection to the engine. This is achieved by an injection pressure adapted to the operating status of 200 to 2,500 bar, a variable start of injection, and the possibility of using pre-injection and secondary-injection events (even highly retarded secondary-injection events are possible). In this way, the common-rail system helps to raise specific power output, lower fuel consumption, reduce noise emission, and decrease pollutant emission in diesel engines.

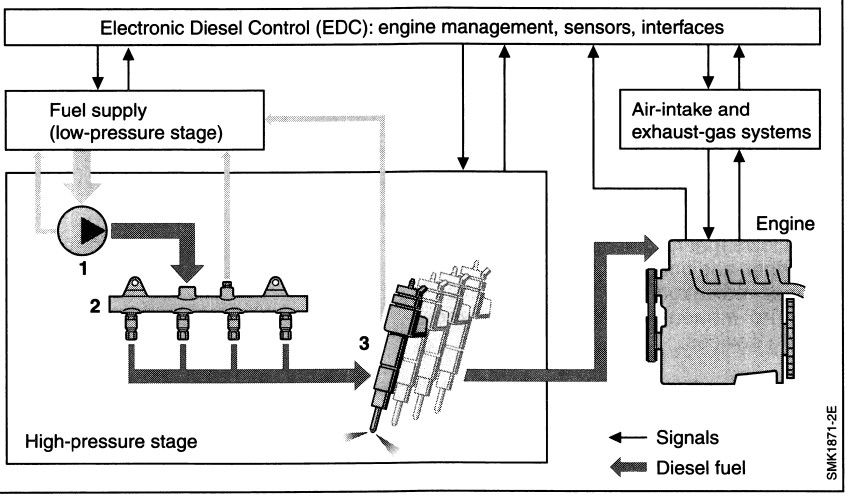
Common rail is today the most commonly used fuel-injection system for modern passenger-car and commercial-vehicle diesel engines.

Design

The common-rail system consists of the following main component groups (Figure 1): the low-pressure system comprising the components of the fuel-supply system, the high-pressure system comprising the high-pressure pump, the rail, the injectors and the high-pressure fuel lines, and Electronic Diesel Control (EDC).

The key components of the common-rail system are the injectors connected to the rail. They are fitted with a rapid-action valve (solenoid valve or piezo actuator) which opens and closes the nozzle. This

Figure 1: System areas of engine management with the common-rail fuel-injection system
1 High-pressure pump, 2 Rail, 3 Injectors.



permits control of the injection process for each cylinder.

Fuel supply

In common-rail systems for passenger cars and light-duty commercial vehicles, electric fuel pumps or gear pumps are used to deliver fuel to the high-pressure pump. It is exclusively gear pumps that are used for heavy-duty commercial vehicles (see Low-pressure fuel supply).

Systems with electric fuel pump

Systems with an electric fuel pump are increasingly gaining acceptance in the passenger-car market. There are several reasons for this:

- Using a controlled electric fuel pump results in a lower power demand (CO_2 advantage).
- As well as supplying the fuel-injection system, an electric fuel pump permits in parallel the operation of suction jet pumps in the tank; these pumps fill the swirl pot and thus ensure the presence of fuel in the intake area (see Fuel-delivery module).
- The fact that the pressure build-up in the low-pressure system is not dependent on engine speed improves the starting response.

The electric fuel pump is usually part of the in-tank unit (in the fuel tank) or occasionally is fitted in the fuel line (in-line). It intakes fuel via a pre-filter and delivers it to the high-pressure pump at a pressure of 4 to 6 bar (Figures 2a and 2c). The maximum delivery rate is, depending on engine performance, 150 to 240 *l*/h.

Systems with gear pump

The gear pump is flanged to the high-pressure pump and is driven by its input shaft (Figure 2b). In this way, the gear pump starts delivery only after the engine has started. Delivery rate is dependent on the engine speed and reaches rates up to 400 *l*/h at pressures up to 7 bar.

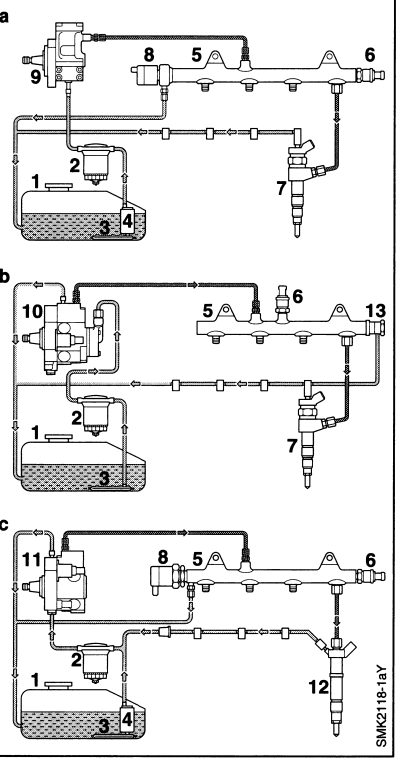
Combination systems

There are also applications where the two pump types are used. The electric fuel pump improves starting response, in

Figure 2: Common-rail systems for passenger cars (examples)

a) Continuous-delivery radial-piston pump: pressure control on the high-pressure side with pressure-control valve.
b) Demand-controlled radial-piston pump: pressure control on the suction side with metering unit flanged to the high-pressure pump.
c) Demand-controlled radial-piston pump: two-actuator system with pressure control on the suction side via the metering unit and pressure control on the high-pressure side via the pressure-control valve.

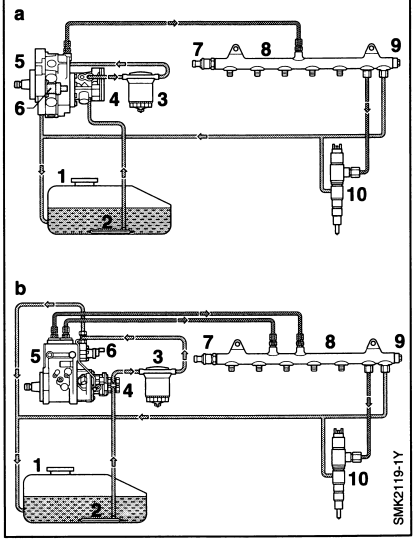
1 Fuel tank, 2 Fuel filter,
3 Preliminary filter, 4 Electric fuel pump,
5 Rail, 6 Rail-pressure sensor,
7 Solenoid-valve injector,
8 Pressure-control valve,
9 Continuous-delivery high-pressure pump,
10 Demand-controlled high-pressure pump with mounted gear presupply pump and metering unit,
11 Demand-controlled high-pressure pump with metering unit,
12 Piezo injector,
13 Pressure limiter.



particular for hot starts, since the delivery rate of the gear pump is lower when the fuel is hot, and therefore, thinner, and at low pump speeds. In addition, the electric fuel pump ensures the operation of suction jet pumps in the tank.

Combination systems are however rarely used, due to the increased costs and the increased power demand for two pumps.

Figure 3: Common-rail systems for commercial vehicles (examples)
a) Demand-controlled radial-piston pump with pressure control on the suction side via the metering unit,
b) Demand-controlled two-plunger in-line pump with pressure control on the suction side via the metering unit.
1 Fuel tank,
2 Preliminary filter,
3 Fuel filter,
4 Gear presupply pump,
5 High-pressure pump,
6 Metering unit,
7 Rail-pressure sensor,
8 Rail,
9 Pressure limiter,
10 Injector.



Fuel filtering
In passenger-car systems with an electric fuel pump the fuel filter is located on the pressure side between the electric fuel pump and the high-pressure pump. In systems with a gear pump the fuel filter is fitted between the fuel tank and the gear pump. In contrast, the fuel filter (fine filter) is fitted on the pressure side in commercial-vehicle systems. For this reason, an exterior fuel inlet is required, in particular when the gear pump is flanged to the high-pressure pump (Figure 3).

Pressure generation
In the common-rail pressure-accumulator fuel-injection system, the functions of pressure generation and fuel injection are separate. The injection pressure is generated independent of the engine speed and the injected fuel quantity. Electronic Diesel Control controls each of the components.

The functions of pressure generation and fuel injection are separated by the rail's accumulator volume. The pressurized fuel is held in the accumulator volume ready for injection.

A continuously operating high-pressure pump driven by the engine produces the desired injection pressure. Pressure in the fuel rail is maintained irrespective of engine speed or injected fuel quantity.

The high-pressure pump is a radial-piston pump. On commercial vehicles, in contrast to passenger cars, an in-line pump is sometimes fitted.

Pressure control
Different procedures are used for pressure control.

Control on the high-pressure side
The desired rail pressure is regulated on the high-pressure side by a pressure-control valve (Figure 2a). Fuel not required for injection flows back to the low-pressure circuit via the pressure-control valve. This type of control loop allows rail pressure to react rapidly to changes in operating point (e.g. in the event of load changes).

Fuel-delivery control on the suction side
Another way of controlling rail pressure is to control fuel delivery on the suction side (Figures 2b and 3). The metering unit flanged on the high-pressure pump makes sure that the pump delivers exactly the right quantity of fuel to the fuel rail in order to maintain the injection pressure required by the system. In a fault situation, the pressure-relief valve prevents rail pressure from exceeding a maximum. Fuel-delivery control on the suction side reduces the quantity of fuel under high pressure and lowers the power input of the pump. This has a positive impact on fuel consumption. At the same time, the temperature of the fuel flowing back to the fuel tank is reduced in contrast to the control method on the high-pressure side.

Two-actuator system
The two-actuator system (Figure 2c) combines pressure control on the suction side via the metering unit and control on the high-pressure side via the pressure-control valve, thus marrying the advantages of suction-side fuel-delivery control with the dynamic response of control on the high-pressure side. Another advantage compared with control on the low-pressure side only is that the high-pressure side is also controllable when the engine is cold. The high-pressure pump then delivers more fuel than is injected and pressure is controlled by the pressure-control valve. Compression heats the fuel, thus eliminating the need for an additional fuel heater.

Fuel injection
The injectors spray fuel directly into the engine's combustion chambers. They are supplied by short high-pressure fuel lines connected to the fuel rail. The engine control unit controls the switching valve integrated into the injector to open and close the injector nozzle.

The injector opening times and system pressure determine the quantity of fuel delivered. At a constant pressure, the fuel quantity delivered is proportional to the switching time of the solenoid valve. This is, therefore, independent of engine and pump speeds (time-based fuel injection).

Hydraulic performance potential
Separating the functions of pressure generation and fuel injection opens up further degrees of freedom in the combustion process compared with conventional fuel-injection systems; the injection pressure is more or less freely selectable within the program map.

The common-rail system allows a further reduction in exhaust-gas emissions by introducing pre-injection events and multiple injection events and also attenuating combustion noise significantly. Multiple injection events of up to seven per injection cycle can be generated by triggering the rapid-action switching valve several times. Hydraulic pressure is used to augment nozzle-needle closing, ensuring rapid termination of the injection process.

Control and regulation
Operating principle
The engine control unit detects the accelerator-pedal position and the current operating states of the engine and vehicle by means of sensors (see "Electronic Diesel Control"). The data collected includes: the crankshaft speed and the crankshaft angle, the rail pressure, the charge-air pressure, the intake-air, coolant and fuel temperature, the inducted air mass, and the wheel speed (for calculating the driving speed). The electronic control unit evaluates the input signals. In sync with combustion, it calculates the triggering signals for the pressure-control valve or the metering unit, the injectors, and the other actuators (e.g. the EGR valve, exhaust-gas turbocharger actuators, etc.).

Basic functions
The basic functions involve the precise control of diesel-fuel injection timing and fuel quantity at the reference pressure. In this way, they ensure that the diesel engine has low consumption and smooth running characteristics.

Correction functions for calculating injection time
A number of correction functions are available to compensate for tolerances between the fuel-injection system and the engine.

Fuel-balancing control
These tolerances result in different torque build-up of individual cylinders. The consequences of these torque differences are irregular engine running and increased exhaust-gas emissions. Injected-fuel-quantity corrections are calculated from the resulting speed fluctuations. Smooth running is increased by specifically adapting the injection time for each cylinder (smooth-running control).

Injector delivery compensation
Injector delivery compensation enables the fuel-quantity deviations of brand-new injectors to be corrected. To this end a wide range of measured data is recorded during injector production for each injector and then applied to the injector in the form of a data matrix code. Information about the lift behavior is also added for a piezo in-line injector. These test data are transmitted during vehicle production to the engine ECU. These values are used during engine operation to compensate deviations in metering and switching performance.

Zero delivery calibration
To achieve both noise reduction and emission targets it is particularly important to be safely in control of small pre-injections over the service life of the vehicle. Injector fuel-quantity drifts must therefore be compensated. To this end a small quantity of fuel is specifically injected into a cylinder in overrun conditions. The speed sensor detects the resulting torque increase as a small dynamic change in speed. This torque increase which is not felt by the driver is clearly linked to the injected fuel quantity. The process is repeated in succession for all the cylinders and for different operating points. A learning algorithm establishes the smallest changes in the pre-injection quantity and corrects the activation duration for the injectors accordingly for all the pre-injections.

Fuel-quantity mean-value adaptation
The deviation of the actually injected fuel quantity from the setpoint value is required to correctly adapt exhaust-gas recirculation and charge-air pressure. Fuel-quantity mean-value adaptation determines the value for the fuel quantity averaged over all the cylinders from the signals from the λ sensor and the air-mass meter. Correction values for the injected fuel quantity are calculated from the comparison of setpoint and actual values.

Supplementary functions
Additional open- and closed-loop control functions perform the tasks of reducing exhaust-gas emissions and fuel consumption, or providing added safety and convenience. Examples of such functions are control of exhaust-gas recirculation, boost-pressure control, cruise control, and the electronic vehicle immobilizer.

The integration of Electronic Diesel Control in a vehicle's overall system network facilitates the exchange of data with, for example, transmission control or air conditioning. A diagnosis interface facilitates the analysis of stored system data when the vehicle is serviced.

Injectors

Solenoid-valve injectors

Design and operating principle

The solenoid-valve injector can be subdivided into a number of function modules: the injection nozzle, the hydraulic servo valve to actuate the valve plunger, the nozzle needle, and the solenoid valve.

The fuel is routed from the high-pressure port (Figure 4a) via an inlet passage to the injection nozzle and via the inlet restrictor into the valve control chamber. The valve control chamber is connected by the outlet restrictor, which can be opened by the solenoid valve, to the fuel return.

The function of the injector can be subdivided into four operating states when the engine and the high-pressure pump are operating. The operating states are caused by the balance of forces acting on the injector components. The nozzle spring closes the nozzle when the engine is not running and there is no pressure in the rail.

All the solenoid-valve injectors have a hydraulic port on the solenoid group for the return flow to the low-pressure system. The return flow is made up of the control quantity (only while the valve is open) and the leakage of the guides at the nozzle, the valve plunger and the solenoid valve (injectors with a maximum pressure greater than 1,600 bar). To increase hydraulic efficiency, injectors with a maximum injection pressure greater than 2,000 bar have no leakage at the nozzle guide and valve plunger.

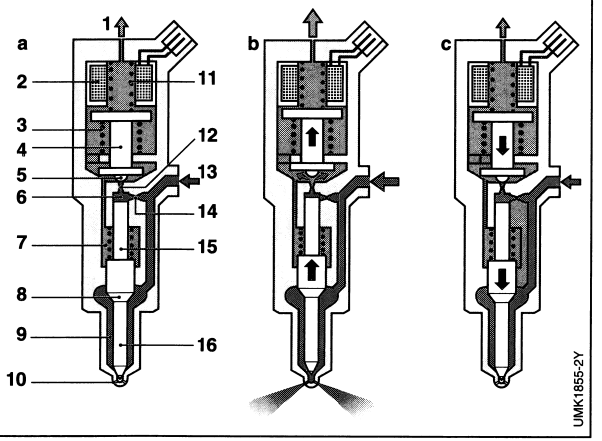
Nozzle closed (rest position)

In its rest position, the solenoid valve is not triggered (Figure 4a). The solenoid valve, in response to the force of the solenoid-valve spring, closes the valve seat, thereby stopping the flow through the outlet restrictor. Inside of the valve control chamber, the pressure rises to the pressure in the fuel rail. The same pressure is also present in the nozzle chamber. The forces acting on the end faces of the control plunger resulting from the fuel-rail pressure and the force of the nozzle spring hold the nozzle needle in the closed position against the opening force acting on the pressure shoulder of the needle.

Figure 4: Functioning principle of solenoid-valve injector (schematic representation)

a) Nozzle closed (rest position),
b) Nozzle opens (start of injection),
c) Nozzle closes (end of injection).

1 Fuel return,
2 Solenoid coil (solenoid valve),
3 Overstroke spring,
4 Solenoid armature,
5 Valve seat, here with pressure-loaded ball valve,
6 Valve control chamber,
7 Nozzle spring,
8 Pressure shoulder of nozzle needle,
9 Chamber volume of injection nozzle,
10 Nozzle-body seat with injection orifices,
11 Solenoid-valve spring,
12 Outlet restrictor,
13 High-pressure port,
14 Inlet restrictor,
15 Valve plunger (control plunger),
16 Injection-nozzle needle.



Nozzle opens (start of injection)
During the opening phase of triggering (Figure 7) the magnetic force of the electromagnet exceeds the spring force of the solenoid-valve spring. The armature opens the valve completely in the ensuing pickup-current phase. This releases the flow through the outlet restrictor (Figure 4b).

The rapid opening of the solenoid valve and the required rapid switching times are achieved by controlling solenoid-valve triggering in the ECU at high voltages and currents (Figure 7). After a short time, the increased pickup current is reduced to a lower holding current.

When the outlet restrictor opens, fuel flows from the valve control chamber to the armature chamber and then via the fuel-return line to the fuel tank. The pressure in the valve control chamber drops. Fuel flows continuously through the inlet restrictor into the control chamber. This prevents the pressure in the control chamber from dropping completely. The flows of the outlet and inlet restrictors are adapted to the dynamic response (switching times) of the solenoid valve. The pressure at the nozzle needle remains at the level of that in the rail. The reduction in pressure in the valve-control chamber reduces the force acting on the control plunger and opens the nozzle needle. Fuel injection commences.

Nozzle open
The rate of movement of the nozzle needle is determined by the difference in the flow rates through the inlet and outlet restrictors. Fuel is injected into the combustion chamber at a pressure approaching that in the fuel rail.

The balance of forces in the injector is similar to that during the opening phase. At a given rail pressure, the fuel quantity injected is proportional to the length of time that the solenoid valve is open. This is entirely independent of the engine or pump speed (time-based injection system).

Nozzle closes (end of injection)
When the solenoid valve is no longer triggered, the solenoid-valve spring presses the armature down, closes the valve seat, and thus stops the flow through the outlet restrictor (Figure 4c). When the outlet restrictor closes, pressure in the control chamber rises again to that in the fuel rail via the inlet restrictor. The higher pressure exerts a greater force on the control plunger. The force on the valve-control chamber and the nozzle-spring force then exceed the force acting on the nozzle needle, and the nozzle needle moves in the direction of the nozzle-body seat. The flow from the inlet restrictor determines the speed with which the nozzle needle closes. The fuel-injection cycle comes to an end when the nozzle needle is resting against its seat, thus closing off the injection orifices.

This indirect method is used to trigger the nozzle needle by means of a hydraulic servo system because the forces required to open the nozzle needle rapidly cannot be generated directly by the solenoid valve. The control quantity required in addition to the injected fuel quantity passes through the valve control chamber restrictors into the fuel return.

Multiple injections of up to eight injection pulses are possible (pilot injection, main injection, secondary injection) within one injection cycle. The minimum possible time interval is approximately 150 µs.

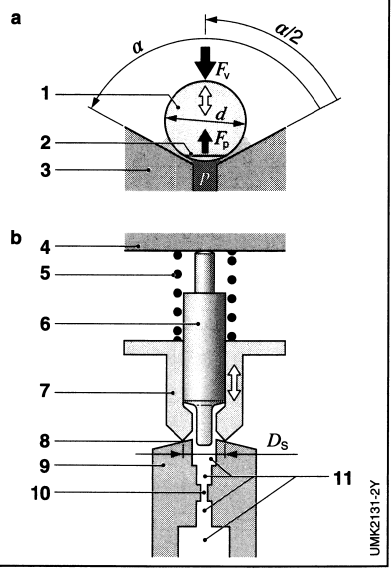
Injector variants
A distinction is made between two different valve concepts with solenoid-valve injectors:
- Injectors with a pressure-loaded ball valve (valve forces act against the applied rail pressure)
- Injectors with a pressure-compensated valve (valve forces are virtually independent of the rail pressure)

Pressure-loaded ball valve
In the case of a pressure-loaded ball valve, the pressure of the compressed fuel acts on the area resulting from valve-seat angle and ball diameter (Figure 5a, see also Figure 4). This pressure generates an opening force. The spring force must be at least of sufficient magnitude that valve remains closed in the non-active state. In practice, the spring force is approximately 15 % greater than the hydraulic force at maximum injection pressure in order on

the one hand to achieve a sufficient dynamic response when the valve closes, and on the other hand to ensure sufficient leak-tightness when closed.

Pressure-compensated valve
In the case of a pressure-compensated valve, there is no area on which the pressure acts and a force can be generated in the opening direction (Figure 5b).

Figure 5: Valves of injector variants
a) Pressure-loaded ball valve,
b) Pressure-compensated valve.
1 Valve ball with diameter d,
2 Valve-seat diameter D, $D = d \sin(90° - \alpha/2)$,
3 Valve section with valve sealing seat,
4 Counter-bearing,
5 Solenoid-valve spring,
6 Armature pin in solenoid assembly (diameter D_S),
7 Solenoid armature,
8 Valve seat (diameter D_S),
9 Valve section with outlet restrictor and valve chamber,
10 Outlet restrictor,
11 Valve chamber.
α Valve-seat angle,
p Rail pressure,
F_p Hydraulic force ($F_p = \pi/4 \cdot D^2 p$),
F_V Valve spring force.



Bosch offers solenoid-valve injectors for passenger-car diesel engines from a maximum rail pressure of more than 1,600 bar with a pressure-compensated valve. The demands placed on the injectors can be met with these valves for modern diesel engines, since a maximum valve opening cross-section can also be attained at pressures in excess of 1,600 bar. This improves the hydraulic stability of the servo valve with which the needle lift is controlled. In addition, the valve lift is also significantly reduced at high rail pressures in order also to achieve the required dynamic response for injection pressures in excess of 1,600 bar and to reduce the sensitivity of the valve dynamic response to external influences. The required minimum time intervals between two injections are achieved on account of the high dynamic response. The power demand for triggering the solenoid valves can also be reduced.

These injector variants with a pressure-compensated valve have a potential for rail pressures in excess of 2,000 bar. There is also the possibility of damping pressure pulsations in the injector with a mini-volume and thereby increasing metering precision in the case of multiple injections.

Further development of solenoid-valve injectors
The further development of solenoid-valve injectors is currently being driven in two directions:
- higher maximum permissible injection pressures,
- improving metering precision and reducing the injected-fuel-quantity tolerances.

Higher injection pressures
By increasing the injection pressure, it is possible, with the same engine power, to improve combustion by reducing the injection-orifice diameter at the nozzle and to lower the untreated emissions that arise in the combustion chamber. Solenoid-valve injectors are currently available for a maximum pressure of 2,500 bar. It is possible to increase the pressure further, but it becomes increasingly more difficult to achieve the degree of strength required on the part of the injector components. The

materials required to achieve this are expensive and the amount of work involved in machining is difficult and extensive.

Metering precision

To improve the metering precision of the injected fuel quantity, sensors are available with which the opening time of the valve and the nozzle needle can be measured and a control loop can be established with software programmed in the engine ECU. Then the opening time of the solenoid valve and consequently the opening time of the nozzle needle can be regulated via the activation period. In this way, it is possible to eliminate to a large extent the deviation of the injected fuel quantity through external influences (e.g., fuel temperatures and viscosity) and through influences in the injector (e.g., wear on the components).

In Bosch solenoid-valve injectors this sensor is integrated in the solenoid of the pressure-compensated valve (Figure 6). The pressure in the valve control chamber is transferred via the armature pin and the pressure distributor to a piezo sensor; the piezo sensor uses the pressure force to generate an electrical signal. This is transmitted via an additional pin in the injector connector to the engine ECU, where it is evaluated.

The pressure characteristic in the control chamber correlates with the opening of the valve and the closing of the nozzle

Figure 6: Solenoid-valve injector with piezo sensor in the solenoid assembly

1 Valve control chamber,
2 Piezo-sensor element,
3 Discharge fitting, 4 Pressure distributor,
5 Electrical connection (three pins),
6 Solenoid sleeve, 7 Solenoid retaining nut,
8 Solenoid core with coil,
9 Mini-rail volume (high pressure),
10 Extrusion coating, 11 Nozzle needle,
12 Valve spring, 13 Armature pin,
14 Diffuser geometry.

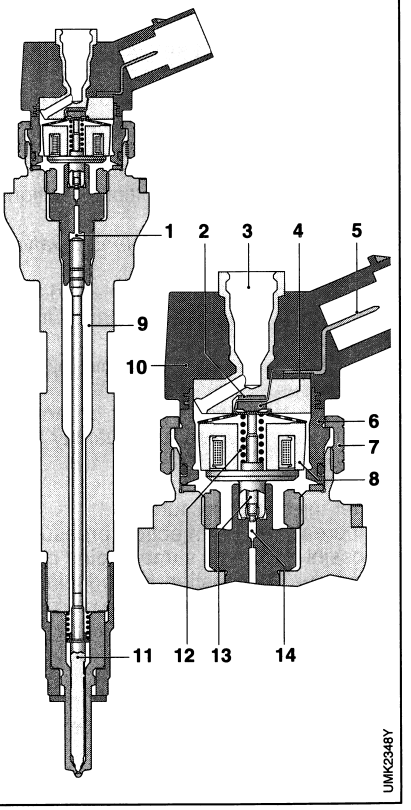
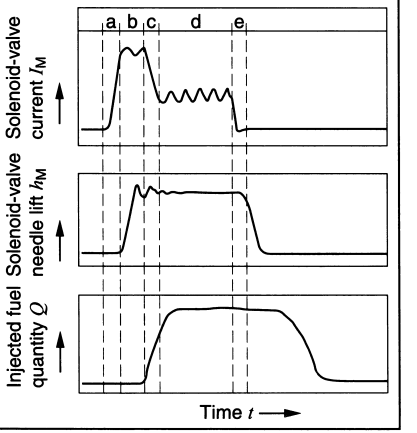


Figure 7: Triggering sequence of a high-pressure solenoid valve for a single injection event

a Opening phase,
b Pickup-current phase,
c Transition to holding-current phase,
d Holding-current phase,
e Switch-off.



needle. In this way, a control loop can be established to regulate the injected fuel quantities.

Triggering the solenoid-valve injector
Triggering of the solenoid valve is divided into five phases (Figure 7; the following details apply to injectors with a pressure-loaded ball valve; small values are sufficient for pressure-compensated valves).

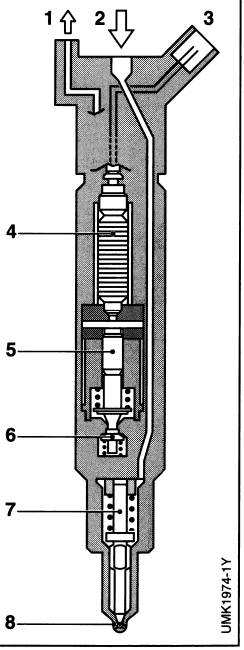
Initially, in order to ensure tight tolerances and high levels of reproducibility for the injected fuel quantity, the current for opening the solenoid valve features a steep, precisely defined flank and increases rapidly up to approx. 20 A (opening phase). This is achieved by means of a booster voltage of up to 50 V which is generated in the ECU and stored in a capacitor (boost-voltage store). When this voltage is applied across the solenoid valve, the current increases several times faster than it does when only battery voltage is used. During the pickup-current phase, battery voltage is applied to the solenoid valve. Current control limits pickup current to approx. 20 A.

In order to reduce power loss in the ECU and injector, the current is dropped to approx. 13 A in the holding-current phase. The energy which becomes available when pickup current is reduced to holding current is routed to the booster-voltage capacitor.

The solenoid valve is closed when the current is switched off, and energy is also released in the process. This energy is also routed to the booster-voltage capacitor. The difference between the energy drawn from the booster-voltage capacitor and the energy returned to it is routed to the booster-voltage capacitor via a step-up chopper integrated in the ECU from the vehicle electrical system. Recharging from the vehicle electrical system is performed until the original voltage level required to open the solenoid valve is reached.

Figure 8: Structural design of piezo inline injector

1 Fuel return,
2 High-pressure port,
3 Electrical connection
4 Piezo actuator,
5 Hydraulic coupler,
6 Servo valve (control valve),
7 Nozzle module with nozzle needle,
8 Injection orifice.



Piezo inline injector

Design and requirements
The design of the piezo inline injector is divided into its main modules in the schematic (Figure 8):
- actuator module (piezo actuator and encapsulation, contact, components for support and force transmission of the actuator),
- hydraulic coupler,
- servo valve (control valve), and
- nozzle module.

A direct response of the nozzle needle to piezo-actuator operation is achieved by coupling the servo valve closely to the nozzle needle. The delay between the electric start of triggering and hydraulic response of the nozzle needle is about 150 µs. At the same time, on account of the high switching force of the piezo actuator, a servo-valve flight time (time between end stops) of just 50 µs and a practically bounce-free switching time are achieved. This meets the contradictory requirements with regard to high nozzle-needle speeds and extremely small reproducible injected fuel quantities.

Similarly to the solenoid-valve injector, a control quantity is terminated via the servo valve to activate an injection. Necessitated by the design, however, the piezo inline injector otherwise contains no leakage points between the high-pressure and low-pressure circuits. The result is an increase in the hydraulic efficiency of the overall system.

Up to eight injection pulses can be set per injection cycle. In this way, the injection can be adapted to the requirements of the respective engine operating point. The maximum possible number of injection pulses decreases in the upper rpm range.

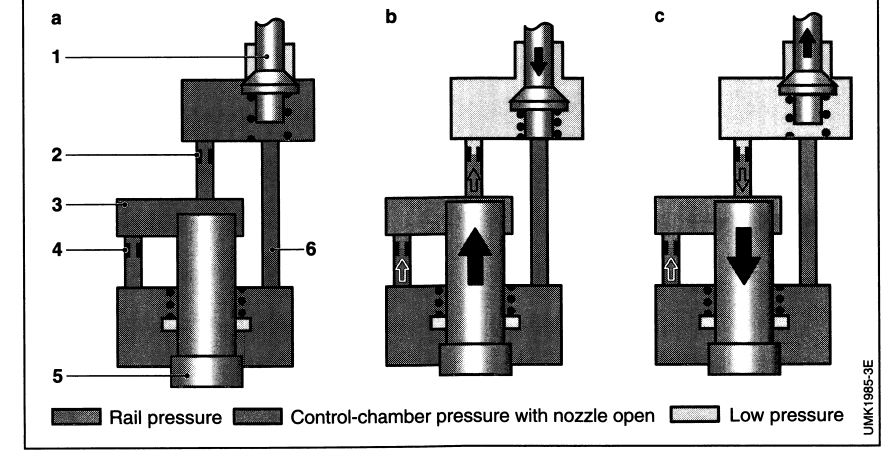
Operating principle of the servo valve

The nozzle needle on a piezo inline injector is controlled indirectly by the servo valve. The desired injected fuel quantity is set, taking into account the prevailing rail pressure, via the activation duration of the valve. In its non-triggered state, the injector is in the starting position and the servo valve is closed (Figure 9a), i.e. the high-pressure section is separated from the low-pressure section. The nozzle is kept closed by the rail pressure exerted in the control chamber.

When the piezo actuator is triggered, it increases in length and this increase is transmitted via the hydraulic coupler to the servo valve. This process opens the servo valve and closes the bypass passage (Figure 9b). The flow-rate ratio between the outlet and inlet restrictors lowers pressure in the control chamber and the nozzle opens. The control quantity flows via the servo valve to the low-pressure circuit of the overall system and from there back to the fuel tank.

To start the closing process, the actuator is discharged, and the servo valve closes to release the bypass passage at the same time (Figure 9c). The control chamber is then refilled by reversing the inlet and outlet restrictors, and the nozzle needle is closed again. The injection process is completed as soon as the nozzle needle reaches the nozzle seat again.

Figure 9: Operating principle of servo valve
a) Start position, b) Nozzle needle opens (bypass closed),
c) Nozzle needle closes (bypass open)
1 Valve pin, 2 Outlet restrictor, 3 Control chamber,
4 Inlet restrictor, 5 Nozzle needle, 6 Bypass.



Operating principle of the hydraulic coupler

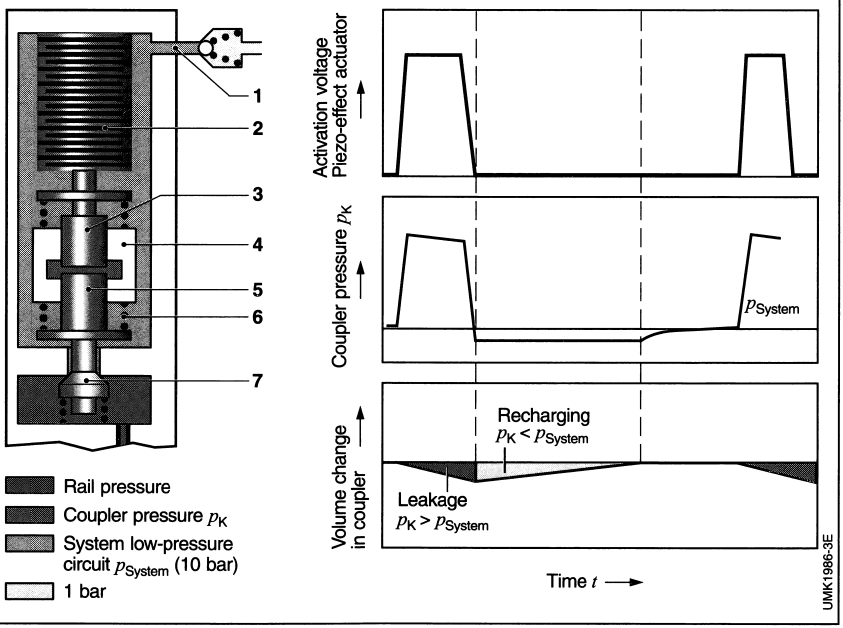
Another vital component in the piezo inline injector is the hydraulic coupler (Figure 9). It provides for compensation of length tolerances of the steel and ceramic components (e.g. caused by the varying thermal expansion of ceramic and steel or by clamping forces on the holder body). On the other hand, it adjusts the translation of actuator stroke and actuator force to the level required on the servo-valve side. The translation ratio is obtained from the diameters of the coupler plunger and valve plunger.

The actuator module and the hydraulic coupler are immersed in the diesel fuel flow, which through the system low-pressure circuit at the injector return is at a pressure of about 10 bar. When the actuator is not triggered, pressure in the hydraulic coupler is in equilibrium with its surroundings, and the coupler does not exert any force on the valve pin. Changes in length caused by temperature or clamping forces acting on the holder body are compensated by small leakages that flow via the guide clearances of the coupler and valve plungers between the coupler gap and the coupler surrounding area. This maintains a coupling of forces between the piezo actuator and the servo valve at all times.

A voltage dependent on the rail pressure (110 to 160 V) is applied to the actuator in order to create an injection event. This causes the pressure in the coupler to increase and a switching force to be exerted on the valve pin. The servo valve opens if this switching force exceeds the closing force on the valve pin caused by the rail pressure. On account of the higher

Figure 10: Operating principle of hydraulic coupler

1 Low-pressure rail with pressure-holding valve (fuel return), 2 Piezo actuator, 3 Coupler plunger, 4 Hydraulic coupler, 5 Valve plunger (lower coupler plunger), 6 Valve-plunger spring, 7 Valve pin.



pressure in the coupler than in the surroundings, a small leakage volume flows through the plunger guide clearances out of the coupler into the low-pressure circuit (10 bar) of the injector. Even when the coupler is repeatedly actuated in quick succession during an engine combustion cycle, there is no significant effect on the injector function when the coupler is drained.

The quantity missing in the hydraulic coupler is refilled in the pauses between the triggering pulses of the piezo actuator. This occurs in the reverse direction via the plunger guide clearance, whereby the valve-plunger spring generates a vacuum pressure of max. 10 bar inside the coupler against the surroundings. The guide clearances and the low-pressure level are matched to refill the hydraulic coupler entirely before the next engine combustion cycle starts.

Further development of piezo injectors

The further development of piezo injectors is being driven forward in two directions:

- higher maximum permissible injection pressures,
- shortening of the necessary idle times between the injections.

High-pressure pumps

Design and requirements

The high-pressure pump is the interface between the low-pressure and high-pressure stages of the common-rail system. Its function is to make sure there is always sufficient fuel under pressure in all engine operating conditions. At the same time, it must operate for the entire service life of the vehicle. This includes providing a fuel reserve that is required for quick engine starting and rapid pressure rise in the fuel rail.

The high-pressure pump generates a constant system pressure for the high-pressure accumulator (fuel rail) independent of fuel injection. For this reason, fuel – compared to conventional fuel-injection systems – is not compressed during the injection process.

3-, 2- and 1-plunger radial-piston pumps are used as the high-pressure pump to generate pressure. In a 3-plunger pump an eccentric shaft provides the stroke motion for the pump plungers, while in 2- and 1-plunger pumps this is provided by a camshaft. 2-plunger in-line fuel-injection pumps are also used in commercial vehicles.

Preferably, the high-pressure pump is fitted to the diesel engine at the same point as a conventional distributor injection pump. The pump is driven by the engine via coupling, gearwheels, chain, or toothed belt. Pump speed is therefore coupled to engine speed via a fixed gear ratio.

High-pressure pumps are used in a number of different designs in passenger cars and commercial vehicles. Within the pump generations there are versions that have different delivery rates (50–550 *l*/h) and delivery pressures (900–2,500 bar).

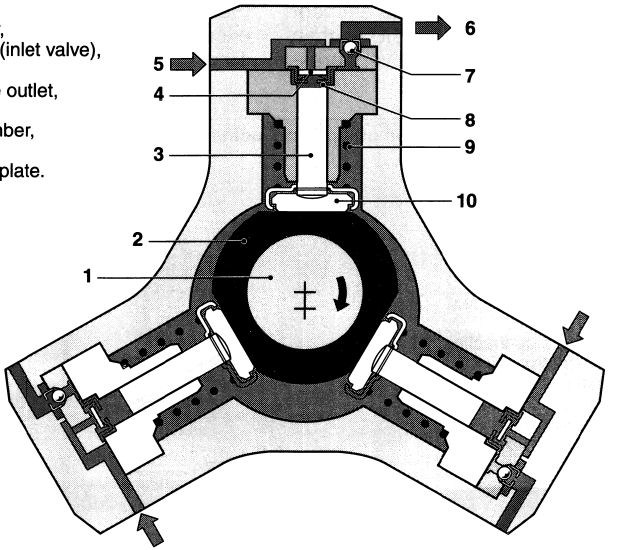
3-plunger radial-piston pump

Design

The drive shaft is centrally mounted in the housing of the high-pressure pump (Figures 11 and 12). The pump elements are arranged radially with respect to the

Figure 11: 3-plunger radial-piston pump (cross-section)

1 Drive shaft with eccentric,
2 Polygon ring,
3 Pump plunger,
4 Suction valve (inlet valve),
5 Fuel supply,
6 High-pressure outlet,
7 Outlet valve,
8 Element chamber,
9 Valve spring,
10 Plunger base plate.



central bearing and offset by 120°. The polygon ring fitted to the drive-shaft eccentric forces the pump plunger to move up and down. Force is transmitted between the eccentric shaft and the pump plunger via the plunger base plate attached to the plunger base.

Fuel delivery and compression

The presupply pump – an electric fuel pump or a mechanically driven gear pump – delivers fuel via a filter and water separator to the inlet of the high-pressure pump. The inlet is located inside the pump on passenger-car systems with a gear pump flanged to the high-pressure pump. An overflow valve is fitted behind the inlet. If the delivery pressure of the presupply pump exceeds the opening pressure (0.5–1.5 bar) of the overflow valve, the fuel is pressed through the restriction bore of the overflow valve into the lubrication and cooling circuit of the high-pressure pump. With its eccentric, the drive shaft moves the pump plunger up and down to mimic the eccentric lift. Fuel passes through the suction valve into the element chamber and the pump plunger moves downward (induction stroke).

When the bottom-dead center of the pump plunger is exceeded, the suction valve closes, and the fuel in the element chamber can no longer escape. It can then be pressurized beyond the delivery pressure of the presupply pump. When the rising pressure exceeds the back pressure from the rail, the outlet valve opens and the compressed fuel passes into the high-pressure circuit. The high-pressure ports of the three pump elements are combined inside the pump housing so that only one high-pressure line runs to the rail.

The pump plunger continues to deliver fuel until it reaches its top-dead center position (delivery stroke). The pressure then drops so that the outlet valve closes. The pump plunger moves downward in response to the force of the valve spring and the remaining fuel in the dead volume is depressurized.

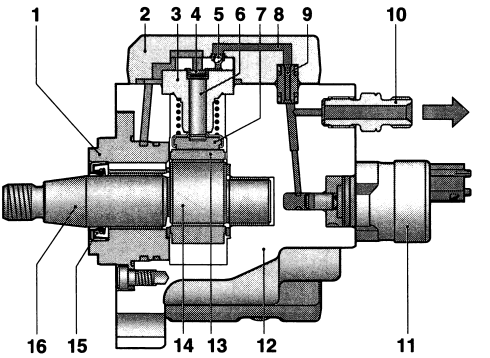
When the pressure in the element chamber drops below the difference between the pre-delivery pressure and the opening pressure of the suction valve, the suction valve reopens and the process starts over.

Transmission ratio

The delivery quantity of a high-pressure pump is proportional to its rotational speed. In turn, the pump speed is dependent on the engine speed. The transmission ratio between the engine and the pump is determined in the process of adapting the fuel-injection system to the engine so as to limit the volume of excess fuel delivered. At the same time, it makes sure that the engine's fuel de-

Figure 12: 3-plunger radial-piston pump (longitudinal section)

1 Flange,
2 Cylinder head,
3 Pump barrel,
4 Suction valve (inlet valve),
5 Outlet valve,
6 Pump plunger,
7 Plunger base plate,
8 High-pressure passage,
9 Connecting piece,
10 High-pressure connection,
11 Pressure-control valve (for constant-delivery high-pressure pump),
12 Pump housing,
13 Polygon ring,
14 Eccentric,
15 Shaft seal,
16 Input shaft.



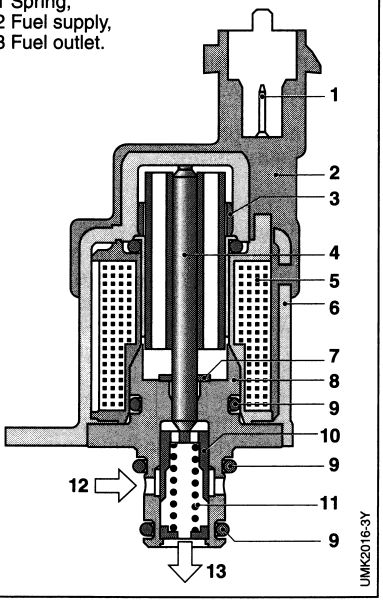
mand at WOT is covered to the full extent. Possible transmission ratios relative to the crankshaft are between 1:2 and 5:6, i.e. the high-pressure pump is geared up. Pump speed is thus higher than engine speed. Higher transmission ratios are required for commercial vehicles on account of the low engine speeds.

Delivery rate

As the high-pressure pump is designed for high delivery quantities, there is a surplus of pressurized fuel when the engine is idling or running in part-load range. This excess fuel is returned in first-generation systems to the fuel tank via the pressure-control valve seated on the rail or flanged to the pump. As the compressed fuel expands, the energy imparted by compression is lost; overall efficiency drops. Compressing and then expanding the fuel also heats the fuel.

Figure 13: Design of metering unit

1 Plug with electrical interface,
2 Solenoid housing,
3 Bearing,
4 Armature with tappet,
5 Solenoid-valve winding with coil body,
6 Bowl,
7 Residual air-gap washer,
8 Magnetic core,
9 O-ring,
10 Plunger with metering slots,
11 Spring,
12 Fuel supply,
13 Fuel outlet.



Demand control

An improvement in energy efficiency is possible by means of fuel-delivery control of the high-pressure pump on the fuel-delivery side (suction side). Fuel flowing into the pump elements is metered by an infinitely variable solenoid valve (metering unit) mounted on the high-pressure pump (Figure 13). This valve adapts the fuel quantity delivered to the rail to system demand via the metering slots in the plunger. The plunger operated by the solenoid valve frees up a metering orifice depending on its position via the metering slots. The solenoid valve is activated by means of a PWM signal (pulse width modulation).

This fuel-delivery control not only drops the performance demand of the high-pressure pump, it also reduces the maximum fuel temperature.

1- and 2-plunger radial-piston pumps

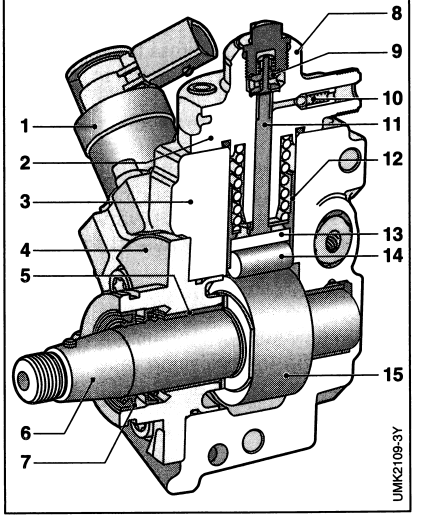
Requirements

The delivery stroke of the pump elements causes pressure pulsations in the rail which give rise to fluctuations in injected fuel quantity in the 3-plunger pumps. In order to comply with the ever more stringent emissions limits, precision of injection coupled with minimal fluctuations in injected fuel quantity is increasingly gaining in importance. 1- and 2-plunger radial-piston pumps facilitate injection-synchronous delivery, i.e. the delivery stroke of the pump elements is synchronous with the induction stroke of the engine cylinders. In this way, the pump delivery for each engine cylinder is always at the same crankshaft angle.

With one or two pump elements, all engines with three to eight cylinders can be injection-synchronously operated by adapting the transmission ratio of 1:2 to 1:1 between engine and pump speeds.

Design
This high-pressure pump is a radial-piston pump of 1- or 2-plunger design. Its components are (Figure 14):
- Aluminum housing, to which only low pressure is applied
- One or two pump elements with high-pressure-resistant steel cylinder heads with integrated high-pressure valve and high-pressure port
- Cam drive assembly with roller tappet, which converts the rotary motion of the camshaft via the cams (dual cam with 180° offset) into a stroke motion of the pump in the cylinder head. The camshaft is guided in the mounting flange and housing in two friction bearings.

Figure 14: 1-plunger radial-piston pump
1 Metering unit,
2 Pump element,
3 Pump housing,
4 Mounting flange,
5 Friction bearing,
6 Input shaft (camshaft),
7 Shaft seal,
8 Cylinder head,
9 Suction valve (inlet valve),
10 High-pressure valve (non-return valve) in high-pressure port (fuel inlet not visible in this illustration),
11 Pump plunger,
12 Roller tappet,
13 Roller support,
14 Drive roller,
15 Dual cam.



The high pressure is generated in the pump element. 1- or 2-plunger pumps are used, depending on the engine's displacement and number of cylinders. Two pump elements are required to cover the fuel demand of larger engines. In the 2-plunger version, the pump elements are arranged in a vee configuration at 90° to each other.

The great overlap length between the cylinder wall and the pump plunger results in low leakage losses when the fuel is compressed. On the other hand, short leakage times result – on account of the high delivery frequency (two strokes per rotation per plunger) and the small dead volume in the cylinder head – in further optimization of efficiency and thus reduced fuel consumption.

The 90° vee configuration of the cylinder heads in the 2-plunger pump means that there is no overlapping of the induction strokes. The charge of the two pump elements is therefore identical (uniformity of delivery).

The high-pressure port is connected to the rail by way of one (in the 1-plunger pump) or two (in the 2-plunger pump) high-pressure lines. The high pressure is not condensed in the housing, but instead routed directly outwards from the cylinder head. The housing therefore does not require any high-pressure- and strength-increasing measures.

Low-pressure circuit
The total fuel delivered by the presupply pump (electric fuel pump or gear pump flanged to the high-pressure pump) is routed through the pump interior to the overflow valve and to the metering unit. The volume of fuel used for lubrication and cooling is thus greater than in previous pumps. The overflow valve regulates the pump internal pressure and thereby protects the housing against overpressure.

The entire low-pressure path is dethrottled on account of the large cross-sections in such a way as to ensure that the pump elements are filled even at high engine speeds. Fuel-quantity metering is

performed on the low-pressure side with the metering unit or the electric suction valve (see Demand control).

High-pressure circuit
The fuel pilot-controlled by the metering unit passes in the induction phase through the suction valve into the element chamber, and is compressed to high pressure during the ensuing delivery phase and delivered by the high-pressure valve and the high-pressure line to the rail.

Demand control with metering unit
The concept of the metering unit used on these pumps corresponds to that of the metering unit used for the 3-plunger radial-piston pump (Figure 13); however, they differ in their design.

Demand control with electric suction valve
The electric suction valve is, like the metering unit, a solenoid valve which is electrically activated. Unlike the metering unit, which is a proportional valve, the electric suction valve is a switching valve.

The electric suction valve is an actively controlled valve (Figure 15) that is seated in place of a conventional suction valve on the injection-pump cylinder head. The plunger chamber is filled with fuel in the high-pressure-pump suction phase. The suction valve is initially kept open during the delivery phase. The delivered fuel is routed back into the low-pressure circuit until the desired fuel quantity is reached in the plunger chamber. The suction valve is then closed and the fuel quantity remaining in the plunger chamber is delivered to the rail. The electric suction valve is activated as a function of the cam-rotation angle.

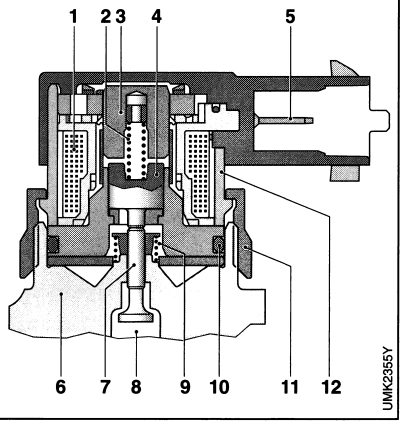
Design and function
The electric suction valve consists of the magnet group and the hydraulic module. The suction valve is open at zero current. In its non-activated state the suction valve is kept open via the armature spring against the hydraulic flow forces. When the suction valve is activated, the solenoid coil is energized to build up a magnetic field. The magnetic field generates a force in the armature which acts against the armature spring. If the magnetic force is greater than the spring force, the armature moves towards the pole body and thereby closes the suction valve.

Features of electric suction valve
The electric suction valve offers more flexibility in activation for high-efficiency engines. It is possible to determine with the cam-angle-dependent control process when the pump begins fuel delivery. This enables precise, dynamic rail-pressure control.

Active opening of the suction valve creates very slight throttling effects during the suction phase, which results in high hydraulic pump efficiency.

A further feature of the electric suction valve to be highlighted is its capacity not to deliver any fuel to the rail despite the pump being driven. This results in the possibility of engine cutout.

Figure 15: Electric suction valve
1 Solenoid coil,
2 Armature spring,
3 Pole body,
4 Armature,
5 Flat connector,
6 Cylinder head of high-pressure pump,
7 Valve plunger,
8 Plunger chamber of high-pressure pump,
9 Compression spring,
10 O-ring,
11 Union nut,
12 Solenoid sleeve.



2-plunger in-line piston pump

Design

This demand-controlled high-pressure pump for rail pressures up to 2,500 bar is only used in commercial vehicles. This is a 2-plunger pump with an in-line design, i.e. the two pump elements are arranged one after the other in relation to the camshaft axial direction (Figure 16). This high-pressure pump comes in both oil-lubricated and fuel-lubricated variants.

A spring seat forms the positive link between the roller tappet and the pump plunger. The rotary motion of the camshaft is converted via the cams into a stroke motion of the pump plungers. The plunger spring ensures pump-plunger return. The combined inlet and outlet valve is located at the top of the pump element.

A gear presupply pump with a high gear ratio is located on the camshaft extension. Its function is to draw fuel via the fuel inlet from the fuel tank and route it through the fuel outlet to the fuel fine filter. From there, the fuel passes through another line to the metering unit located on the upper section of the high-pressure pump.

Lube oil is supplied either directly via the mounting flange of the pump or via a side-mounted inlet. The lube oil is returned to the engine oil pan.

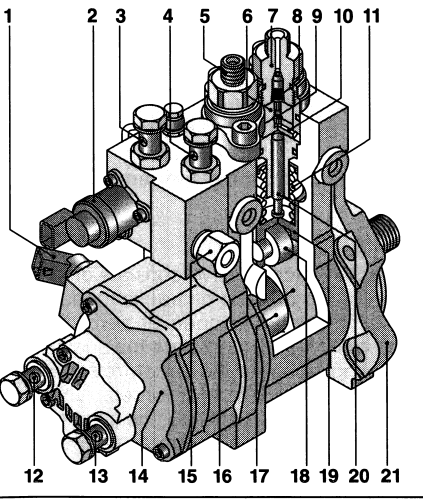
Operating principle

When the pump plunger moves from top dead center towards bottom dead center, the suction valve opens on account of the fuel pressure (pre-delivery pressure). The fuel is drawn into the element chamber in response to the downward movement of the pump plunger. The outlet valve is closed by the valve spring.

As the pump plunger moves upwards, the suction valve closes and the enclosed fuel is compressed. When the rail pressure is exceeded, the outlet valve opens and the fuel is delivered through the high-pressure port to the rail. This increases the pressure in the rail. The rail-pressure sensor measures the pressure, from which the engine ECU calculates the triggering signals (PWM) for the metering unit. The metering unit regulates the fuel quantity provided for compression according to the current demand.

Figure 16: 2-plunger in-line piston pump

1 Speed sensor (pump speed),
2 Metering unit,
3 Fuel supply for metering unit (from fuel filter),
4 Fuel return line to fuel tank,
5 High-pressure port,
6 Valve body,
7 Valve holder,
8 Outlet valve with valve spring,
9 Suction valve (inlet valve) with valve spring,
10 Fuel supply to pump element,
11 Plunger spring,
12 Fuel supply from fuel tank,
13 Fuel outlet to fuel filter,
14 Gear presupply pump,
15 Overflow valve,
16 Concave cam,
17 Camshaft,
18 Roller bolt with roller,
19 Roller tappet,
20 Pump plunger,
21 Mounting flange.



Rail

Function

The function of the rail is to maintain the fuel at high pressure. In so doing, the accumulator volume dampens pressure fluctuations caused by fuel pulses delivered by the pump and the fuel-injection cycles. This ensures that, when an injector opens, the injection pressure remains constant. On the one hand, the accumulator volume must be large enough to meet this requirement. On the other hand, it must be small enough to ensure a fast enough pressure rise on engine start.

Besides acting as a fuel accumulator, the fuel rail also distributes fuel to the injectors.

Application

The tubular rail can vary depending on the various constraints which determine the way it is fitted to the engine. It has connections for fitting the rail-pressure sensor and the pressure limiter or the pressure-control valve (Figure 17).

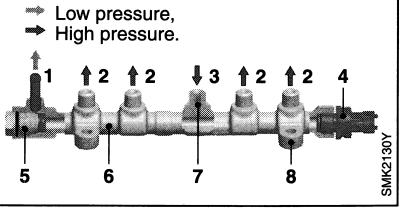
The pressurized fuel delivered by the high-pressure pump passes via one or two high-pressure fuel lines to the rail inlet. From there, it is distributed via high-pressure lines to the individual injectors. The cavity inside the fuel rail is permanently filled with pressurized fuel during engine operation.

The fuel pressure is controlled by means of Electronic Diesel Control (EDC), whereby the fuel pressure is measured by the rail-pressure sensor and – depending on the system – regulated by demand control or via the pressure-control valve to the desired value. The pressure limiter is used as an alternative to the pressure-control valve – depending on system requirements – and its function is to limit fuel pressure in the fuel rail to the maximum permissible pressure in the event of a fault.

Some rail types make use of restrictors in the inlet and outlet which additionally dampen pressure pulsations caused by pump delivery and injections. If fuel is now drawn from the rail for an injection, the pressure is the rail remains virtually constant.

Figure 17: Design of a rail with pressure limiter

1 Fuel return (low pressure),
2 High-pressure ports injectors,
3 High-pressure port to high-pressure pump (one or two ports),
4 Rail-pressure sensor.
5 Pressure limiter,
6 Rail body,
7 Restrictor (pressed-in, optional),
8 Mounting lug (engine mounting).



Rail types

Two different rail types are used: the forged rail (hot-forged rail) and the welded rail (laser-welded rail). Both types are used by Bosch, the hot-forged rail being the preferred variant. The welded rail is used in series production up to 2,000 bar and is not subject to further development.

In the case of a hot-forged rail, the blank for machining is produced by forging from bar stock. The internal geometry and the rail interfaces of the rail body are produced by deep-hole boring, drilling and milling. A corrosion-resistant surface is then applied. Finally, the add-on components are mounted and functionally tested.

The forging process affords more possibilities in the shaping of the external geometry than laser welding. One advantage is the possibility of shaping the external geometry with regard to weight optimization.

The forged rail is used for passenger-car and commercial-vehicle applications in series production up to 2,500 bar, and a further pressure increase is planned for the next generations.

Time-controlled single-cylinder pump systems

Unit-injector system for passenger cars

System requirements

The electronic controlled unit injector is a single-cylinder pump injection system with integrated high-pressure pump and injection nozzle (Figure 1). The high-pressure line required in other injection systems between the injection pump and the injection nozzle is omitted. This means that the system has particularly good hydraulic performance.

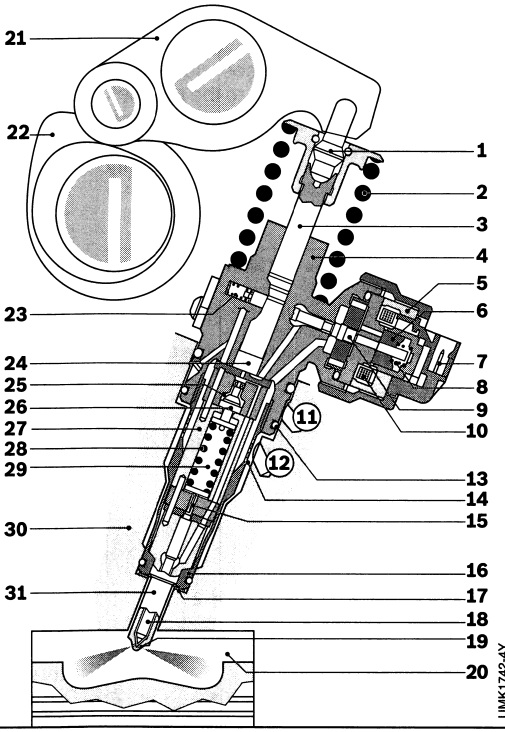
The unit injector is installed in the cylinder head between the valves, with the nozzle protruding into the combustion chamber. It is operated by rocker arms driven by the overhead engine camshaft. Each engine cylinder is allocated a separate unit injector. The start of injection and injection duration are calculated by an electronic control unit and controlled by the high-pressure solenoid valve mounted externally on the pump body.

The unit-injector system (UIS) is designed to meet the demands of modern direct-injection diesel engines with high levels of power density. It is characterized by its compact design, high injection pressures of up to 2,200 bar at full load, and mechanical-hydraulic pilot injection throughout the entire program-map range to substantially reduce combustion noise. The compact design enables a very low high pressure volume with correspondingly high hydraulic efficiency.

Unit-injector systems are no longer used for new developments.

Figure 1: Design of unit injector for passenger cars

1 Ball pin, 2 Return spring,
3 Pump plunger, 4 Pump body,
5 Magnetic core of high-pressure solenoid valve, 6 Armature,
7 Electrical connection,
8 Compensating spring,
9 Solenoid-valve needle,
10 Electromagnet coil,
11 Fuel return (low-pressure stage),
12 Fuel feed, 13 Seal,
14 Fuel inlet passages (approx. 350 laser-drilled holes as filter),
15 Hydraulic stop (damping unit),
16 Retaining nut,
17 Sealing washer,
18 Nozzle needle,
19 Nozzle-needle seat,
20 Engine combustion chamber,
21 Roller rocker arm,
22 Engine camshaft with actuating cam,
23 Solenoid-valve spring,
24 High-pressure chamber (element chamber),
25 Accumulator chamber,
26 Accumulator plunger (bypass plunger),
27 Spring retainer,
28 Nozzle spring (compression spring),
29 Spring-retainer chamber,
30 Engine cylinder head,
31 Integrated injection nozzle.



Operating principle

Intake stroke

The unit injector is filled during the intake stroke of the upward-moving pump plunger. Fuel flows from the low-pressure stage of the fuel supply through the inlet holes into the unit injector. The fuel enters the high-pressure chamber when the solenoid-valve seat is opened.

Prestroke

The rotation of the actuating cam causes the pump plunger to move downwards. When the solenoid valve is still open, the fuel is forced back by the pump plunger from the high-pressure chamber into the low-pressure stage of the fuel supply. Heat is also dissipated from the unit injector (cooling) with the fuel as it flows back.

Delivery stroke and pilot injection

The electronic control unit energizes the coil of the solenoid valve at a particular point in time so that the solenoid-valve needle is pressed onto the solenoid-valve seat and the connection between high-pressure chamber and low-pressure stage is closed (beginning of injection period, BIP). The downward movement of the pump plunger increases the fuel pressure in the high-pressure chamber. The nozzle-opening pressure for pilot injection is approximately 180 bar. When this pressure is reached, the nozzle needle is raised and pilot injection begins. In this phase the nozzle-needle lift is hydraulically limited by a damping unit located between nozzle needle and nozzle spring so that the low injected fuel quantity required can be accurately metered. The accumulator plunger remains on its seat for the time being because the nozzle needle opens first due to its greater hydraulically effective area on which the pressure acts.

As the pressure increases, the accumulator plunger (bypass plunger) is forced downwards and now lifts off its seat. A connection is established between high-pressure chamber and accumulator chamber. The resulting pressure drop in the high-pressure chamber, the increased pressure in the accumulator chamber and the simultaneous increase in the pre-tension of the nozzle spring cause the nozzle needle to close. Pilot injection is terminated. However, the accumulator plunger does not return to its starting position because in its open state it offers the fuel pressure a greater working area than the nozzle needle.

Main injection

The continuing movement of the pump plunger causes the pressure in the high-pressure chamber to increase further. The nozzle-opening pressure for main injection is at approximately 300 bar higher than pilot injection. There are two reasons for this. Firstly, the excursion of the accumulator plunger increases the pre-tension of the nozzle spring. Secondly, due to the bypassing of the accumulator plunger fuel must be forced from the spring-retainer chamber via a restrictor into the low-pressure stage of the fuel supply so that the fuel in the spring-retainer chamber is compressed to a greater extent (pressure-backing). The pressure-backing level is derived from the size of the restrictor in the spring retainer and can be varied. It is therefore possible to achieve a sensible compromise between a low opening pressure for pilot injection (for noise reasons) and as high an opening pressure as possible for main injection, especially at part load (reducing emissions). The stroke and shaft diameter of the bypass plunger determine the length of the interval, the so-called injection interval, between the end of pilot injection and the start of main injection.

When the nozzle-opening pressure is reached, the nozzle needle is raised and fuel is injected into the combustion chamber (actual start of injection). The pressure during the entire injection process continues to rise due to the high delivery rate of the pump plunger.

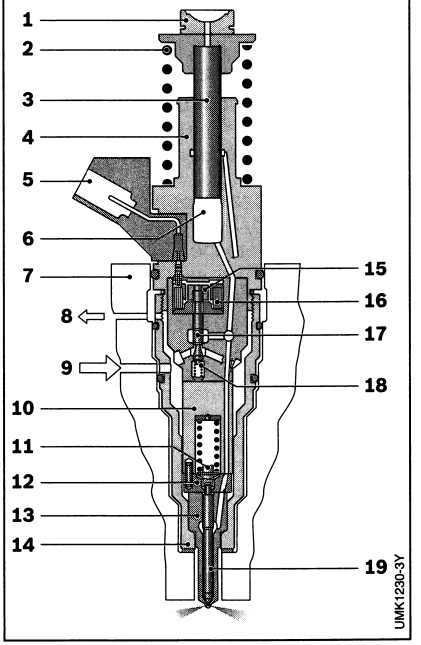
The current flow through the solenoid-valve coil is deactivated to terminate main injection. The solenoid valve opens and releases the connection between high-pressure chamber and low-pressure stage. The pressure collapses. When the nozzle-closing pressure is undershot, the injection nozzle closes and the injection process ends. At this point the accumulator plunger also returns to its starting position. The remaining fuel is returned to the low-pressure stage during the continuing

downward movement of the pump plunger (residual stroke). Heat is also dissipated from the unit injector in the process.

Electronic control allows the values for the start of injection and the injected fuel quantity to be selected as desired from those stored in the program map. This feature, together with the high injection pressures, makes it possible to achieve high power densities combined with low emission levels and low fuel consumption.

Figure 2: Design of unit injector for commercial vehicles

1 Sliding disk, 2 Return spring,
3 Pump plunger, 4 Pump body,
5 Electrical connection,
6 High-pressure chamber (element chamber),
7 Engine cylinder head,
8 Fuel return (low-pressure stage),
9 Fuel feed, 10 Spring retainer,
11 Pressure pin, 12 Intermediate ring,
13 Integrated injection nozzle,
14 Retaining nut, 15 Armature,
16 Electromagnet coil,
17 Solenoid-valve needle,
18 Solenoid-valve spring,
19 Nozzle needle.



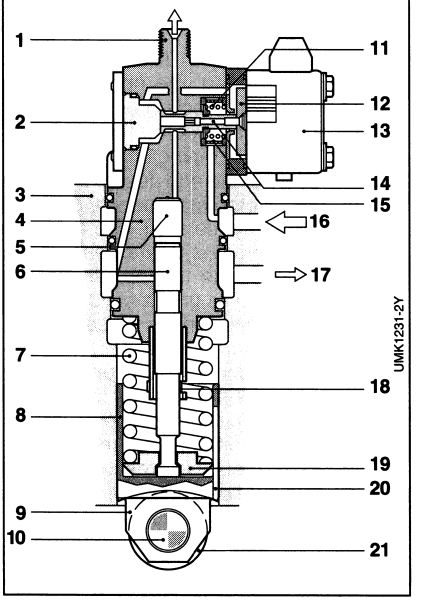
Unit-injector system for commercial vehicles

The unit-injector system for commercial vehicles is largely the same as the system for passenger cars. Due to the larger dimensions for the commercial-vehicle system, the solenoid valve can be integrated in the unit-injector here (Figure 2).

With regard to main injection, the commercial-vehicle system has the same operating principle as the passenger-car system. The systems differ with regard to pilot injection. In the unit-injector System

Figure 3: Design of unit pump

1 Hydraulic high-pressure port,
2 Solenoid-valve needle-travel stop,
3 Engine block, 4 Pump body,
5 High-pressure chamber,
6 Pump plunger, 7 Tappet spring,
8 Roller-tappet shell,
9 Roller tappet, 10 Roller-tappet pin,
11 Solenoid-valve spring,
12 Armature plate,
13 Solenoid-valve housing,
14 Solenoid-valve needle, 15 Filter,
16 Fuel feed, 17 Fuel return,
18 Pump-plunger retainer,
19 Spring seat, 20 Locating groove,
21 Tappet roller.



for commercial vehicles this can be electronically controlled in the lower speed and load range. This significantly reduces combustion noise and improves cold-starting performance.

Unit-pump system for commercial vehicles

The unit-pump system (UPS) is also a modular, time-controlled single-cylinder pump injection system, and is closely related to the unit-injector system. It is used in commercial-vehicle engines and in large engines. Each of the engine's cylinders is supplied by a separate module consisting of a high-pressure plug-in pump with integral, high-speed solenoid valve, a short high-pressure line, and an injection nozzle (Figures 3 and 4). This system is therefore also called Pump-Line-Nozzle (PLN). It permits injection pressures of up to 2,100 bar.

The separation of high-pressure generation and injection permits a simpler mounting on the engine. The hydraulic performance is nevertheless very good due to the shortest line possible. The unit pump is mounted on the side of the engine block. It is operated directly by an injection cam on the engine's camshaft via a roller tappet. The injection nozzle is installed with a nozzle holder in the cylinder head. The lines are made of high-strength, seamless steel tubes. They must be of equal length for the individual pumps of an engine.

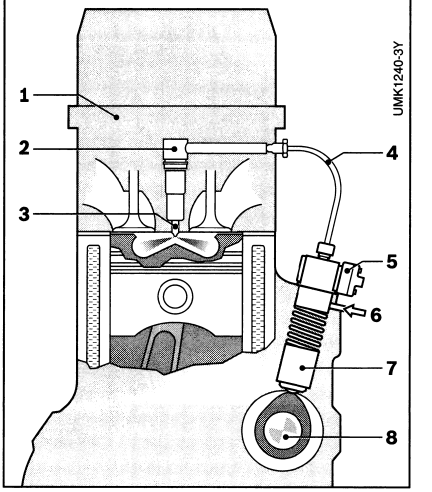
The method of solenoid-valve actuation is the same as that of the unit-injector system. When the solenoid valve is open, the fuel can be drawn into the pump barrel during the pump plunger's intake stroke, and return during the delivery stroke. Only when the solenoid valve is energized, and thus closed, can pressure build up in the high-pressure system between the pump plunger and the nozzle during the pump-plunger delivery stroke. Fuel is injected into the combustion chamber of the engine once the nozzle-opening pressure is exceeded.

Electronic control

The solenoid valves are actuated by an electronic control unit. The ECU analyzes all of the relevant status parameters in the system relative to the engine and its environment, and defines the start of injection and injected fuel quantity for the operating state of the engine at any given time. The start of injection is also controlled by a BIP signal (beginning of injection period) in order to balance out the tolerances in the overall system. Start of injection is synchronized with engine piston position by analysis of the signals from an incremental trigger wheel.

In addition to the basic fuel-injection functions, there are further functions for improving driving smoothness (e.g. surge dampers, idle-speed governors, adaptive cylinder equalization). Diagnosis of the fuel-injection system and the engine is also included in the range of functions. The ECU communicates with other electronic vehicle components (e.g. antilock braking system (ABS), the traction control system (TCS) or transmission-shift control system) via a CAN data bus.

Figure 4: Unit-pump system
1 Engine, 2 Nozzle holder,
3 Injection nozzle, 4 High-pressure line,
5 Solenoid valve, 6 Fuel feed,
7 High-pressure pump (unit pump),
8 Camshaft.



Diesel distributor injection pumps

Diesel distributor injection pumps were used in large numbers from 1962 until after 2000. These pumps are used in 3-, 4-, 5- and 6-cylinder diesel engines in passenger cars, tractors and light and medium commercial vehicles. They generate up to 50 kW per cylinder, depending on engine speed and combustion system. Distributor injection pumps for direct-injection (DI) engines achieve a peak injection pressure of up to 1,950 bar in the nozzle at engine speeds up to 4,500 rpm.

A distinction is made between distributor injection pumps with mechanical control, and those with an electronic governor available in port-controlled versions with a rotary-magnet actuator and versions with solenoid-valve open-loop control.

In passenger cars and commercial vehicles, distributor injection pumps have been superseded by common-rail systems.

Axial-piston distributor pumps

Body

Fuel-supply pump

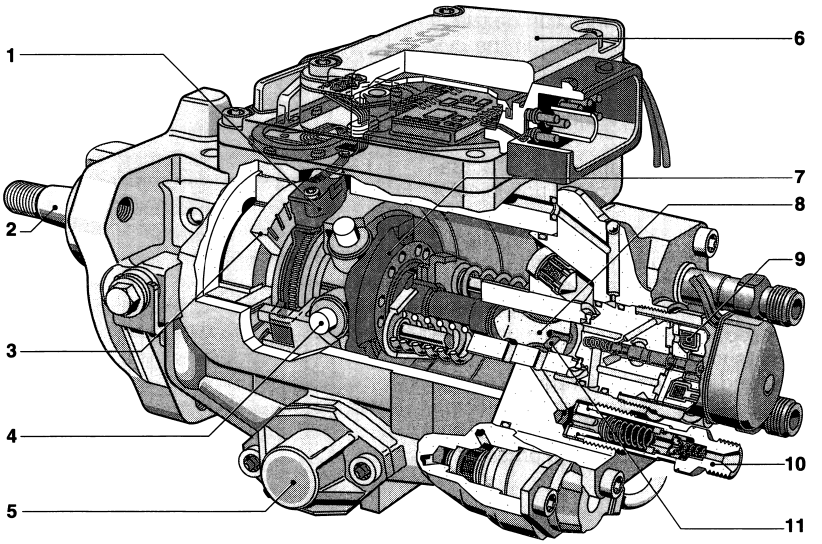
If no presupply pump is fitted in the fuel-injection system, this integral vane-type supply pump draws fuel from the tank and, together with a pressure-control valve, generates an internal pump pressure which increases with engine speed.

High-pressure pump

The axial-piston distributor pump (VE pump) incorporates only one pump element for all cylinders. The distributor plunger displaces the fuel during its axial stroke and rotates at the same time to distribute the fuel through passages to the individual high-pressure connections (Figure 1).

Figure 1: Solenoid-valve-controlled axial-piston distributor pump

1 Angle-of-rotation sensor, 2 Drive shaft, 3 Support ring of vane-type supply pump, 4 Roller ring, 5 Timing device, 6 Pump control unit, 7 Cam plate, 8 Distributor plunger, 9 High-pressure solenoid valve, 10 High-pressure connection, 11 Distributor bore.



A clutch unit transfers the rotation of the drive shaft to the cam plate and the distributor plunger fixed to it. When doing so, the claws of the drive shaft and the cam plate grip into the yoke positioned in between. The cam lifts on the underside of the cam plate turn against the rollers of the roller ring. As a result, the cam plate and distributor plunger perform a stroke movement in addition to the rotation. During each rotation of the drive shaft, the distributor plunger completes a number of strokes equal to the number of engine cylinders to be supplied.

The pump delivers fuel for as long as the spill port in the distributor plunger remains closed off during the working stroke. Delivery ends when the spill port is uncovered by the control collar (Figure 2).

Electronic control for distributor injection pumps with rotary-magnet actuator

In contrast to the mechanically controlled distributor injection pump, the pump with rotary-magnet actuator has an electronic governor and an electronically controlled timing device (Figure 2).

Electronic governor

An eccentrically mounted ball head provides the connection between the distributor injection pump control collar and the rotary-magnet actuator. The actuator's rotary setting determines the position of the control collar, and with it the effective stroke of the pump. A non-contacting position sensor is connected to the rotary-magnet actuator.

The control unit receives various signals from the sensors: accelerator-pedal position; engine speed; air, coolant and fuel temperature; charge-air pressure; atmospheric pressure; etc. It uses these input variables to determine the correct injected fuel quantity, which is then converted to a specific control-collar position with the aid of program maps stored in the unit's memory. The control unit varies the excitation current to the rotary-magnet actuator until it receives a signal indicating convergence between the set-point and actual values for control-collar position.

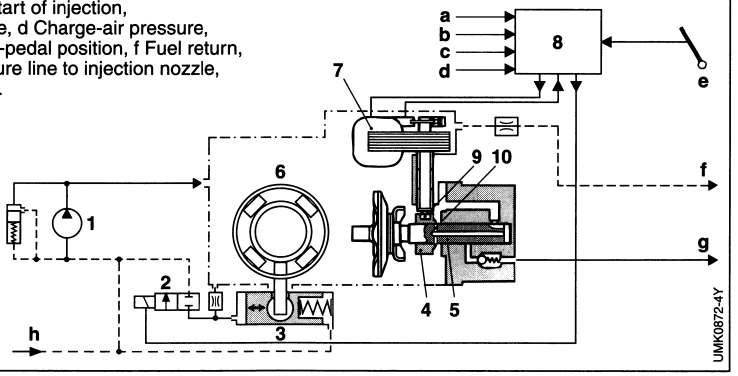
Electronically controlled timing device

The hydraulic timing device with the timing-device solenoid valve rotates the roller ring, depending on the load state and engine speed in such a way that the

Figure 2: Electronic control for axial-piston distributor pumps with rotary-magnet actuator

1 Vane-type supply pump, 2 Timing-device solenoid valve, 3 Timing-device plunger,
4 Control collar, 5 Distributor plunger, 6 Roller ring (view turned through 90°),
7 Rotary-magnet actuator with position sensor, 8 ECU,
9 Eccentrically mounted ball head, 10 Spill port.

Input and output variables:
a Speed, b Start of injection,
c Temperature, d Charge-air pressure,
e Accelerator-pedal position, f Fuel return,
g High-pressure line to injection nozzle,
h Fuel supply.



start of delivery – in relation to the engine piston position – is advanced or retarded. The signal from a sensor in the nozzle-holder assembly, which indicates when the nozzle begins to open, is compared with a programmed setpoint value. The timing-device solenoid valve connected to the working chamber of the plunger in the timing device varies the pressure above the timing-device plunger and thus the position of the timing device. The actuation clock ratio of the timing-device solenoid valve is varied until the setpoint and actual values agree.

Electronic control for solenoid-valve-controlled distributor injection pumps

In the case of solenoid-valve-controlled distributor injection pumps (Figure 1), the fuel is metered by a high-pressure solenoid valve which directly closes off the pump's element chamber. This permits even greater flexibility in fuel metering and for varying the start of injection. The main assemblies are the high-pressure solenoid valve, the ECU, and the incremental angle/time system for angle/time control of the solenoid valve using an angle-of-rotation sensor integrated in the pump.

The solenoid valve closes to define the start of delivery, which then continues until the valve opens. The injected fuel quantity is determined by the length of time the valve remains closed. Solenoid-valve control permits rapid opening and closing of the element chamber independently of engine speed. In contrast to mechanically governed pumps and pumps with a rotary-magnet actuator, direct triggering by means of solenoid valves results in lower dead volumes, improved high-pressure sealing, and therefore greater efficiency.

The fuel-injection pump is equipped with its own, integral control unit for precise start-of-delivery control and fuel metering. Individual pump program maps and example-specific calibration data are stored in this ECU.

The engine control unit determines the start of injection and delivery on the basis of engine operating parameters, and sends this data to the pump control unit via the data bus. The system can control both the start of injection and the start of delivery.

The pump control unit also receives the injected fuel quantity signal via the data bus. This signal is generated by the engine control unit according to the accelerator-pedal signal and other parameters for torque demand. In the pump control unit, the injected fuel quantity signal and the pump speed for a given start of delivery are taken as the input variables for the pump map on which the corresponding actuation period is stored as degrees of cam rotation.

And finally, the actuation of the high-pressure solenoid valve and the desired period of actuation are determined on the basis of the angle-of-rotation sensor integrated in the distributor injection pump. This sensor is used for angle/time control. It consists of a magnetoresistive sensor and a reluctor ring divided into 3° increments interrupted by a reference mark for each cylinder. The sensor determines the precise angle of cam rotation at which the solenoid valve opens and closes. This requires the pump control unit to convert timing data to angular position data and vice versa.

The low fuel-delivery rates at the start of injection, which result from the design of the distributor injection pump, are further reduced by the use of a two-spring nozzle holder. With a hot engine, these low delivery rates permit low basic noise levels.

Pilot injection

Pilot injection allows the combustion noise to be further reduced without sacrificing the objects of the system's design which aim at generating maximum power output at the rated-power operating point. Pilot injection does not require additional hardware. Within a matter of milliseconds, the electronic control unit actuates the solenoid valve twice in rapid succession. For the first injection, only a small amount of fuel is discharged to condition the combustion chamber. The solenoid valve controls the injected fuel quantity with a high degree of precision and dynamic response. Typical pilot-injection fuel quantities are 1.5 mm^3.

Radial-piston distributor pumps

Body

High-pressure pump

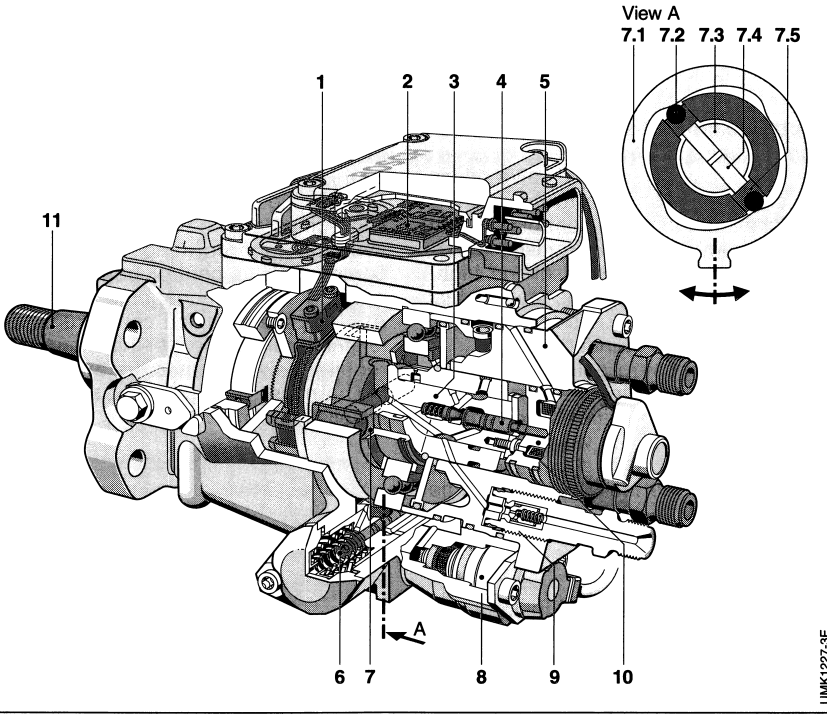
The radial-piston high-pressure pump (VR pump, Figure 3) is driven directly by the distributor-pump drive shaft. It comprises the cam ring, the roller supports and rollers, the delivery plunger, the drive plate, and the front section (head) of the distributor shaft.

The drive shaft drives the drive plate by means of radially positioned guide slots. The guide slots simultaneously act as the locating slots for the roller supports. The roller supports and the rollers held by them run around the inner cam profile of the cam ring that surrounds the drive shaft. The number of cams corresponds to the number of cylinders in the engine.

The drive plate drives the distributor shaft. The head of the distributor shaft holds the delivery plungers which are aligned radially to the drive-shaft axis (hence the name "radial-piston distributor pump").

Figure 3: Solenoid-valve-controlled radial-piston distributor pump

1 Sensor (angle/time system), 2 Pump control unit, 3 Distributor shaft, 4 Solenoid-valve needle, 5 Valve body, 6 Timing device, 7 Radial-piston pump, 7.1 Cam ring, 7.2 Roller, 7.3 Distributor shaft, 7.4 Delivery plunger, 7.5 Roller support, 8 Timing-device solenoid valve, 9 Delivery valve, 10 High-pressure solenoid valve, 11 Drive shaft.



The delivery plungers rest against the roller supports. As the roller supports are forced outwards by centrifugal forces, the deliver plungers follow the profile of the cam ring and describe a cyclical-reciprocating motion.

When the delivery plungers are pushed inwards by the cams, the volume in the central plunger chamber between the delivery plungers is reduced. This compresses and pumps the fuel when the solenoid valve is closed. The fuel is directed through passages in the distributor shaft at defined times to the appropriate outlet delivery valves.

As the cam-drive design employs a direct, positive link, flexibility and compliance remain minimal, so the performance potential is greater. Fuel delivery is shared between at least two radial plungers. The small inertial forces involved mean that steep (fast) cam profiles are possible. The fuel-delivery rate can be further increased by increasing the number of delivery plungers.

Radial-piston distributor pumps for direct-injection engines achieve element-chamber pressures up to 1,100 bar and pressures in the nozzle up to 1,950 bar.

Electronic control system

High-pressure solenoid valve

The high-pressure solenoid valve opens and closes in response to the triggering signals of the pump control unit. The length of time it remains closed determines the delivery period of the high-pressure pump. This means that the fuel quantity can be very precisely metered for each individual cylinder.

The high-pressure solenoid valve is controlled by regulating the current. The pump control unit identifies the contact of the valve needle in the valve seat using the current curve. This allows the exact point at which fuel delivery starts to be calculated and the start of injection to be controlled very precisely.

Timing devices

The hydraulically assisted timing device rotates the cam ring in such a way that the start of delivery – in relation to the engine piston position – is advanced or retarded. The interaction between the high-pressure solenoid valve and the timing device thus varies the start of injection and the injection pattern to suit the operating status of the engine.

The cam ring engages in a cross-slot in the timing-device plunger by means of an adjuster lug so that the axial movement of the timing-device plunger causes the cam ring to rotate. In the center of the timing-device plunger is a control collar which opens and closes the control ports in the timing-device plunger. In axial alignment there is a spring-loaded hydraulic control plunger which defines the required position for the control collar. Under the control of the pump control unit, the timing-device solenoid valve modulates the pressure acting on the control plunger.

This hydraulically assisted timing device can apply higher displacement forces than the hydraulic timing device of the axial-piston distributor pump.

The timing-device solenoid valve acts as a variable throttle. It can vary continuously the control pressure so that the control plunger can assume any position between the fully advanced and fully retarded positions.

Version with full electronic circuitry on injection pump

The latest generation of distributor injection pumps are compact, self-contained systems incorporating an electronic control unit to control both the pump and the engine-management functions. As a separate engine control unit is no longer required, the fuel-injection system requires fewer connectors and the wiring harness is less complex, thus making installation simpler.

Fuel-injection system

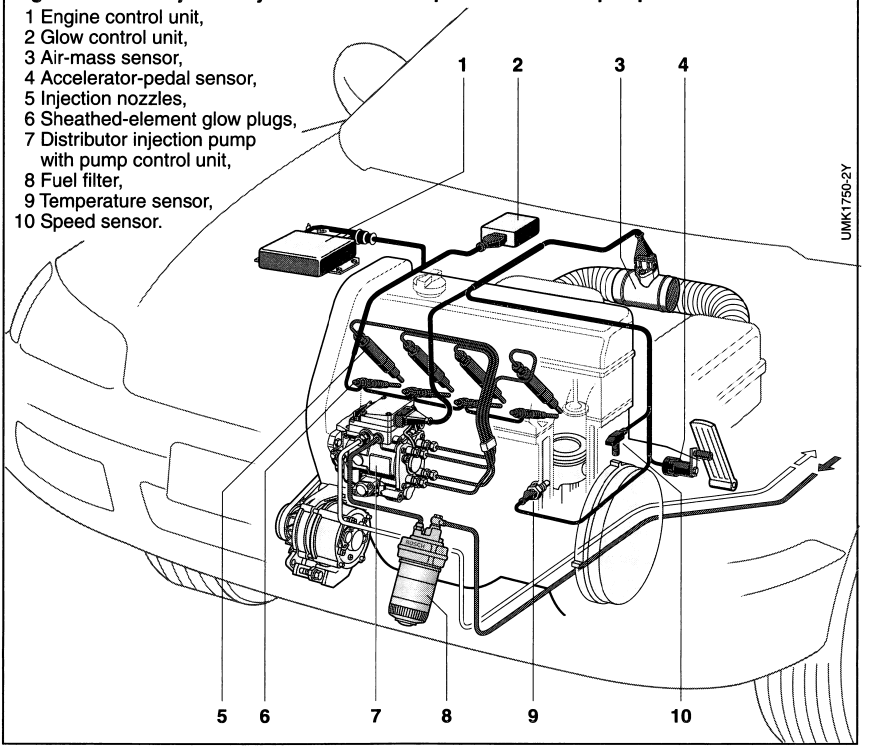
The injection pump is a component part of the fuel-injection system (Figure 4). A diesel fuel-injection system comprises the fuel-supply system (low-pressure stage), the high-pressure component, the injection components and the control system. The fuel-supply system accumulates and filters the fuel. If necessary, an additional fuel pump is present. The injection pump with the lines is the high-pressure component. It generates the high pressure and distributes it at the right time to the corresponding engine cylinders.

In distributor-pump injection systems, the injection component comprises the injection nozzles and the nozzle holders. These nozzle-holder assemblies come in a wide range of types. One nozzle-holder assembly is used on each cylinder. They are mounted with brackets or with hollow screws in the cylinder head. The functions of the injection nozzles are to inject metered quantities of fuel and to prepare the fuel, to shape the injection pattern, and to provide a seal against the combustion chamber. Each injection nozzle consists of the nozzle body with several injection orifices (up to 0.12 mm) and the nozzle needle. The needle is guided in the nozzle-body guide bore to ensure correct alignment between the injection orifices (which are at different angles in the nozzle body) and the engine combustion chamber.

The mechanical or electronic control system for the distributor injection pump is mounted on the pump itself. Some systems have a separate engine control unit. Versions with electronic control are equipped with numerous sensors and setpoint generators.

Figure 4: Diesel injection system with a radial-piston distributor pump

1 Engine control unit,
2 Glow control unit,
3 Air-mass sensor,
4 Accelerator-pedal sensor,
5 Injection nozzles,
6 Sheathed-element glow plugs,
7 Distributor injection pump with pump control unit,
8 Fuel filter,
9 Temperature sensor,
10 Speed sensor.



Start-assist systems for diesel engines

Preheating systems for passenger cars and light commercial vehicles

Hot diesel engines start spontaneously at low outside temperatures. Here, the spontaneous ignition temperature for diesel fuel of 250 °C is achieved with the engine turning at starting speed by compression. Cold diesel engines require, depending on their compression ratio, start assistance with glow plugs at low ambient temperatures (Figure 2).

The use of preheating systems extends far beyond pure start assistance. Thus, in driving mode the glow plugs can be consistently activated when the engine is hot in order to optimize combustion in all operating states.

Design and function of a preheating system

Preheating systems (Figure 1) consist essentially of glow plugs (GLP), the glow control unit (GCU), and preheating software in the engine management. The low-voltage preheating systems mostly used today require glow plugs with nominal voltages below the vehicle system voltage whose heating output is adapted via the electronic glow control unit to the engine requirement.

In DI engines the glow plug extends into the main combustion chamber of the engine cylinder. The air/fuel mixture is directed past the hot tip of the glow plug. In combination with the heating of the intake air during the compression cycle,

Figure 2: Startability of a diesel engine (without preheating) as a function of compression ratio ε

1 ε = 18,
2 ε = 16,
3 ε = 14.

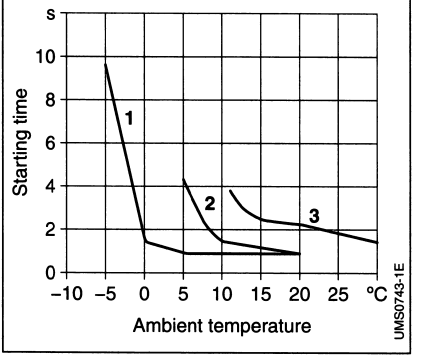
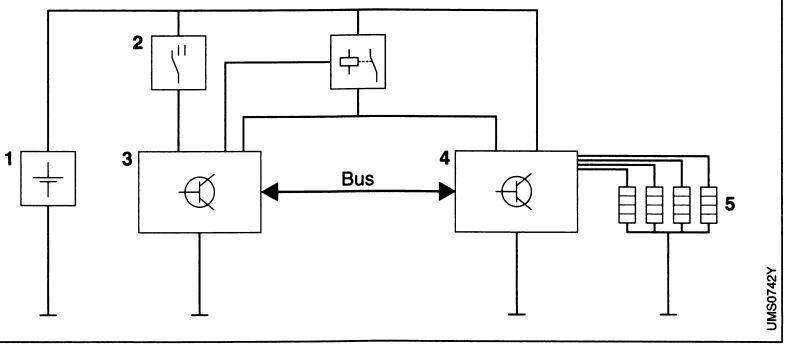


Figure 1: Preheating system

1 Battery, 2 Ignition/starting switch, 3 Engine ECU,
4 Glow control unit, 5 Glow plugs.



the ignition temperature is reached and the air/fuel mixture ignites.

The preheating software in the engine ECU starts and ends the preheating process as a function of the parameters stored in the software (e.g., ignition/starting switch on/off, coolant temperature). The glow control unit activates the glow plugs as directed by the engine ECU with vehicle system voltage via an electronic switch (transistor) during the preheating phases (pre-, standby, start, and post-heating). The supply voltage to the glow plugs is set with pulse-width-modulated activation to values below the vehicle system voltage. Direct operation at vehicle system voltage would destroy the glow element on account of the lower nominal voltage.

Preheating phases

- During preheating the glow plug is heated to operating temperature.
- During standby heating the preheating system maintains a temperature required for starting for a defined period.
- Start heating is applied while the engine is being revved up.
- The post-heating phase begins after starter release.
- Intermediate heating is activated after engine cooling due to coasting or to support particulate-filter regeneration.

In order to achieve as quickly as possible the heating temperature required for engine starting during pre-heating, the glow plugs are briefly operated in this phase with push voltage, which is above the glow-plug nominal voltage. The activating voltage is then reduced to the glow-plug nominal voltage during standby heating.

During start heating, the activation voltage is increased again in order to compensate for the cooling of the glow plug by the cold intake air. This is also possible in the post- and intermediate-heating phases. The required voltage is taken from a number of program maps that are adapted to each engine. The program maps contain among others the parameters engine speed, injected fuel quantity, time after starter release, and coolant temperature.

Map-controlled activation reliably prevents thermal overloading of the glow plugs in all engine operating states. The heating function implemented in the engine ECU contains an overheating-protection facility in the event of repeat heating.

These low-voltage preheating systems facilitate a rapid start similar to a gasoline engine down to approximately –30 °C.

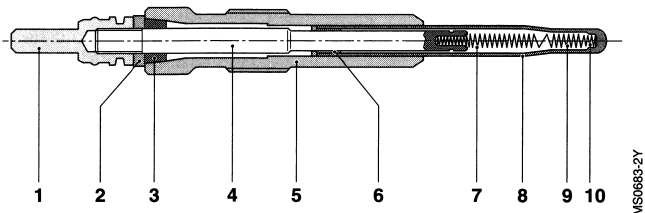
Low-voltage metal glow plug

Design and characteristics

The main component in the sheathed-element glow plug is the tubular heating element (Figure 3). The tubular heating element consists of a hot-gas and corrosion-resistant element sheath which encloses a filament surrounded by compressed magnesium-oxide powder. This filament is made up of two resistors con-

Figure 3: Glow plug

1 Plug connector,
2 Insulator shim,
3 Double seal,
4 Terminal stud,
5 Housing,
6 Element seal,
7 Control filament,
8 Glow tube,
9 Magnesium-oxide powder (filling powder),
10 Heating filament.



nected in series – the heating filament located in the tip of the sheath and the control filament.

Whereas the heating filament has an electrical resistance that is independent of temperature, the control filament has a positive temperature coefficient (PTC). Its resistance increases as the temperature rises.

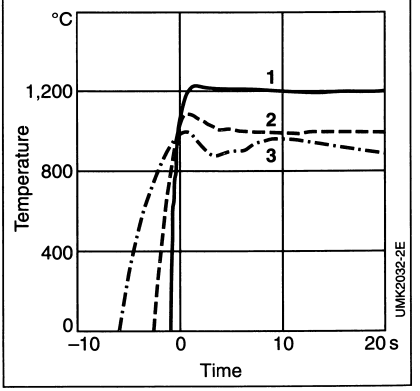
The heating filament is welded into the cap of the element sheath for grounding. The control filament is contacted at the terminal stud, which establishes the connection to the vehicle electrical system.

The glow element is tapered at the front in order to accommodate the heating filament closer to the element sheath. This allows preheating rates of up to 1,000 °C within 3 s with the push mode (Figure 4). The maximum heating temperature is in excess of 1,000 °C. The temperature during start-standby heating and in post-heating mode is approx. 980 °C.

Glow plugs with 11 V nominal voltage are practically no longer used in passenger cars due to their longer heat-up time.

Figure 4: Comparison of preheating curves, from t = 0 s with a flow velocity of 11 m/s

1 Low-voltage ceramic glow plug (e.g., 7 V nominal voltage),
2 Low-voltage metal glow plug (e.g., 5 V nominal voltage),
3 Metal glow plug (11 V nominal voltage).



Operation
When voltage is applied to the glow plug, most of the electrical energy of the heating filament is initially converted into heat. The temperature at the tip of the glow plug increases sharply. The temperature of the control filament – and with it also the impedance – increase with a time delay. The current draw and thus the total heating output of the glow plug decrease and the temperature approaches a steady-state condition. By controlling the operating voltage the temperature is kept below the critical level for the glow plug. Consequently, it can remain in operation for several minutes after the engine has started. This post-heating function results in a more effective engine cold-idle phase with substantially lower noise and emission output.

Reduced emissions in diesel engines with a low compression ratio

By lowering the compression ratio in modern diesel engines from $\varepsilon = 18$ to $\varepsilon = 16$, it is possible to reduce the NO_x and soot emissions while simultaneously increasing the specific power. However, the cold-starting and cold-idling performance is more difficult to control in these engines due to the lower compression. Temperatures at the glow plug of up to 1,250 °C are required in order to achieve minimum exhaust-gas opacity values and smooth running during cold starting and cold idling of these engines. Optimum combustion can only be achieved during the cold-running phase by post-heating lasting several minutes. These requirements can barely be satisfied with metal glow plugs.

Low-voltage ceramic glow plugs

Low-voltage ceramic glow plugs have glow elements made of a ceramic material that is highly resistant to temperatures. Ceramic heating-conductor tracks are embedded into the ceramic material. Their function is comparable to the filaments of metal glow plugs.

Due to their very high resistance to oxidation and thermal shocks, ceramic glow plugs enable an immediate start, as well as minute-long post-heating and inter-

mediate heating at up to 1,250 °C. They are designed for a nominal voltage of 7 V.

Glow control unit
The glow control unit sets the temperature at the glow plug stipulated by the engine management for the respective operating state. The supply voltage is pulse-width-modulated via a transistor for this purpose.

The glow control unit furthermore assumes functions like ground-offset compensation and protection against polarity reversal.

Applications
Low-voltage preheating systems require an application of the system matched to the respective engine in order to set the best temperature in each operating state and simultaneously protect the glow plug against being damaged by overvoltage. The respective activation voltage is stored in program maps.

Regulated preheating system
The regulated preheating system was developed to minimize the application outlay. Ceramic glow plugs exhibit within certain limits a linear behavior between their temperature and their resistance. This behavior is used by an algorithm to calculate in the ECU the current temperature actually applied at the glow element and to automatically adjust the desired temperature accordingly.

Flame start systems for commercial-vehicle diesel engines

Requirements

Commercial-vehicle diesel engines can start down to temperatures of around −20 °C with start assistance. The start durations are dependent on the compression ratio ε and the fuel grade used (cetane number CN, Figure 5). To prevent misfiring and emissions of unburnt hydrocarbons (HC), it is however possible, depending on the compression ratio and the fuel grade used, to utilize already from a coolant temperature of 10 °C air preheating as start assistance (flame start system or grid heater).

Air preheating provides for safe, quick starting down to low outside temperatures. The battery and the starter are protected and fuel consumption is reduced in the starting and warm-up phases. At higher temperatures air preheating would not be necessary for safe starting, but does prevent the formation of white smoke (HC emission, Figure 6).

Figure 5: Starting times without start assistance as a function of compression ratio ε and cetane number CN

1 Starting-time curve for ε = 16.5 and CN = 41,
2 Starting-time curve for ε = 16.5 and CN = 45,
3 Starting-time curve for ε = 16.5 and CN = 51,
4 Starting-time curve for ε = 17.0 and CN = 51,
5 Starting-time curve for ε = 17.75 and CN = 51.

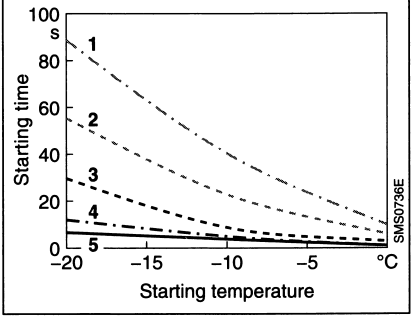
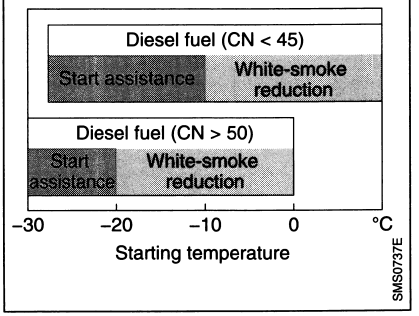


Figure 6: Use of start assistance for starting and to reduce HC emissions



Design

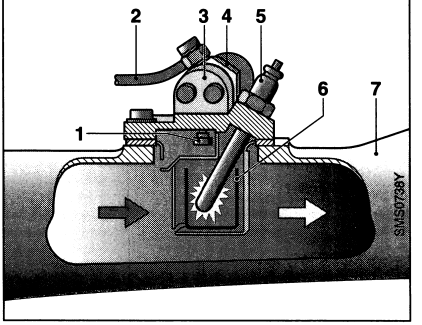
There are two different types of flame start system: the flame start system with burner chamber and the flame glow plug.

Flame start system with burner chamber

The flame start system with burner chamber (Figure 7) consists of a nozzle holder on which a solenoid valve, a nozzle and a rod-type glow plug are mounted. The nozzle holder is attached together with gaskets and the burner chamber by three screws to the charge-air housing.

Figure 7: Flame start system with burner chamber

1 Nozzle, 2 Fuel line,
3 Nozzle holder, 4 Solenoid valve,
5 Rod-type glow plug,
6 Burner chamber,
7 Charge-air housing.

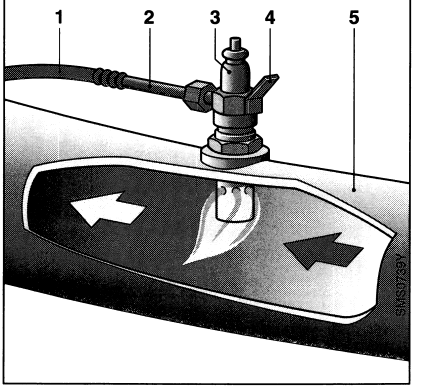


The fuel is drawn after the fuel-system fuel filter and routed to the solenoid valve via a fuel line. The solenoid valve controls the fuel flow while the flame start system is in operation. The glow plug is supplied with electrical power by the flame-start control unit.

Flame glow plug
In contrast to the flame start system with burner chamber, the solenoid valve in the case of the flame glow plug is not mounted directly on the flame start system, but instead is connected by a fuel line to the flame glow plug. The flame glow plug also has a much more compact design (Figure 8).

The glow plug integrated in the flame glow plug is enclosed in a vaporizer tube in the upper area and extends up to the protective tube at the end of the flame glow plug (Figure 9). The fuel is drawn after the fuel filter, routed through the separately mounted solenoid valve, and supplied to the flame glow plug via the fuel line. The fuel enters the flame glow plug via the metering device – to control the fuel quantity. The flame-start control unit assumes the job of controlling the flame start system (current supply to the glow plug and switching on the solenoid valve).

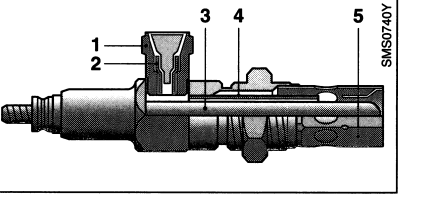
Figure 8: Installed flame glow plug
1 Fuel line,
2 Throttle line,
3 Flame glow plug,
4 Ground connection,
5 Charge-air tube.



Use of flame start systems
The flame start system with burner chamber is used primarily in commercial-vehicle engines with a displacement $V_H > 10\,l$, while the flame glow plug is used in engines with a displacement V_H of 4 to 10 l. The flame start system is installed in the engine charge-air tube in such a way that the heated intake air is uniformly distributed to all the cylinders.

Operating principle
The flame start system is activated below the switch-on temperature defined in the control unit. When the ignition key is turned to the Drive position, the "Flame start system" indicator lamp lights up and the glow plug is energized and preheated. The indicator lamp goes out after a preheating time dependent on the vehicle system voltage (approx. 25 s). The flame start system is operational and the engine can be started. At the outset of the starting operation the solenoid valve opens and the fuel flows to the hot glow plug, where it ignites. The flame that is now produced heats the intake air flowing past. The flame start system is operated after the engine start in the post-flaming time until a stable burn is guaranteed and shuts down according to the parameters defined in the control unit. The parameters for post-flaming are dependent on the coolant temperature.

Figure 9: Flame glow plug (Beru system)
1 Fuel-supply connection,
2 Metering device,
3 Glow plug,
4 Vaporizer tube,
5 Protective tube.



The flame start system is also shut down if the engine is started before the indicator lamp goes out or is not started within 30 seconds of the indicator lamp going out. In this case the driver can put the flame start system into operation again by turning the ignition key back and then turning into to the drive position again.

Control of the glow plug

Glow plug for the flame start system with burner chamber

The glow plug is preheated depending on the vehicle system voltage until a temperature of approximately 1,050 °C is reached. This is followed by the beginning of the period of start readiness (duration 30 s), during which the glow plug is cyclically energized. The cycle ratio is set depending on the current vehicle system voltage in such a way that the glow-plug temperature remains between 1,020 °C and 1,080 °C. With the glow plug in this temperature range the fuel is safely ignited and the glow plug itself is not exposed to thermal overload.

As soon as the engine is started, the solenoid valve opens, the fuel is directed through the nozzle and sprayed onto the spark-plug grooves, where it ignites. In this phase the cycle ratio of the glow plug is increased by the control unit so that the glow plug does not cool in response to the heat dissipated from the vaporizing fuel. This higher cycle ratio is maintained after starting for a further period of around 30 s until all the components are heated and the flame is safely burning. The glow plug is now additionally heated by the flame. The control unit decreases the cycle ratio again to protect against overheating. After post-flaming the glow plug is deactivated and monitored until it has cooled down entirely. Start and cold-running support by the flame start system is concluded.

Flame glow plug

This glow plug is preheated in exactly the same way before starting, depending on the applied vehicle power supply. The preheating times are slightly longer than for the glow plug with burner chamber, since the vaporizer tube is additionally heated during preheating (see Figure 9). When the engine is started, the solenoid valve is opened, fuel flows via the fuel-supply connection and the metering device into the flame glow plug, strikes the hot glow plug, vaporizes between the glow plug and the vaporizer tube, and mixes with the air at the evaporator-tube outlet. The air/fuel mixture ignites at the hot glow plug. The flame produced emerges from the protective tube and heats the intake air flowing past.

After preheating the flame glow plug is likewise cyclically energized. The cycle ratio is decreased in the start-readiness period (30 s) and increased again when the engine is started. In this phase the cycle ratio for the flame glow plug is higher than for the flame start system with burner chamber, since the fuel is vaporized via the supplied electrical energy and not via the burning flame.

Cooling-down phase

The glow plug of the flame start system is monitored in the cooling-down phase. If preheating is performed again in this phase, the control unit calculates the preheating time from a program map depending on the cooling-down time and the voltage in such a way that the glow plug reached the temperature range described above.

Because the glow plugs have heating and control filaments (see Preheating systems for passenger-car diesel engines), the optimum preheating time is set as a function of the ambient conditions.

Grid heater for commercial-vehicle diesel engines

The grid heater is used as a further cold-start aid, particularly for engines which are started under aggravated starting conditions such as constant load when starting (e.g. for hydraulic output) or at altitudes above 2,000 m (e.g. snow groomers).

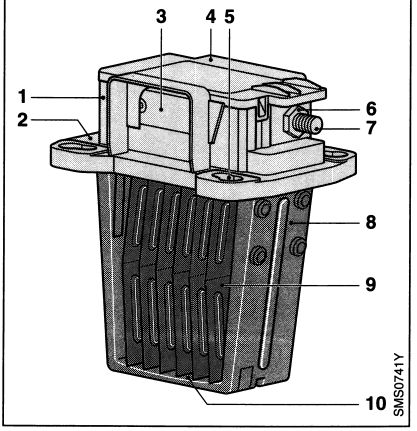
Design

The grid heater (Figure 10) consists of an electric heater element which is installed in the charge-air tube and powered by a relay. A separate ECU or the engine ECU switches on the grid heater on the basis of the stored parameters.

There are different versions which are used, depending on the engine displacement V_H and the vehicle system voltage (12 V or 24 V). Grid heaters with a power input of 1.8 kW are used for engines with a displacement V_H of 4 to 10 l, while grid heaters with a power input of 2.7 kW are used for engines with $V_H > 10\,l$.

Figure 10: Grid heater

1 Housing,
2 Mounting flange,
3 Cover plate,
4 Cover,
5 Mounting bore,
6 Lock nut,
7 Connecting-thread bolt,
8 Retaining frame,
9 Heating filament,
10 Insulating plate.



Operating principle

Depending on the coolant and charge-air temperatures, the grid heater is preheated for periods varying in duration (preheating time t_v = 2 to 28 s) and remains active in the post-heating phase until stable engine running is ensured. The switch-on temperature is 5 °C for engines with V_H = 4 to 10 l and −4° for engines with $V_H > 10\,l$.

Once the preheating time has elapsed, the indicator lamp goes out and the engine can be started. On starting the grid heater is switched off for approximately 2 s so as to make full battery capacity available to the starter. For start and idle support the grid heater is then switched back on and remains in operation until it is switched off after the applied post-heating time.

The grid heater remains in start readiness for only around 5 s (as opposed to 30 s for the flame start system) in order to conserve battery capacity.

Advantages and disadvantages of the grid heater

The grid heater offers the following advantages over the flame start system:

- The intake air is heated before starting
- There are shorter starting times thanks to good ignition conditions from the first revolution
- No burning of oxygen is required to heat the intake air

On the down side, the grid heater requires greater battery capacity.

Exhaust-gas treatment

Catalytic oxidation

The oldest catalytic method of exhaust-gas treatment is aimed at the oxidation of imperfectly combusted fuels. Carbon monoxide (CO) and imperfectly combusted hydrocarbons ($C_xH_yO_z$) – which also include soot and hydrogen (H_2) – are converted in this process with the oxidizing agents present in the exhaust gas, i.e. oxygen (O_2) in particular. These reactions correspond to the stoichiometries of the equations (A) through (D):

(A) $C_x + x\,O_2 \rightarrow x\,CO_2$,
(B) $2\,CO + O_2 \rightarrow 2\,CO_2$,
(C) $2\,H_2 + O_2 \rightarrow 2\,H_2O$,
(D) $4\,C_xH_yO_z + (4x+y-2z)\,O_2 \rightarrow 4x\,CO_2 + 2y\,H_2O$.

As a rule noble metals, their mixtures or their alloys in the form of nanoparticles serve as the catalytically active component. Platinum and palladium are frequently used.

A prerequisite for the most complete sequence of oxidation reactions as possible is an at least stoichiometric quantity of oxygen in the exhaust-gas flow. A catalytic converter can however passivate with high oxygen excess and low exhaust-gas temperatures. The elimination of imperfectly combusted fuels from the exhaust gas works particularly effectively when the catalytic converters and the engine application are matched to each other. The following conditions should be in place for this purpose:

- The engine exhibits low untreated emission of imperfectly combusted fuel, particularly in the cold-start phase.
- The engine application provides the catalytic converter with a balanced air/fuel ratio (lambda, λ).
- The oxidation catalytic converter is heated as quickly as possible by means of heating measures.
- The catalytic converter has a quantity of catalytically active and storing components matched to the emissions and exhaust-gas temperatures.
- The type and quantity of all the components are matched in such a way that the catalytic converter resists to a sufficient extent the loads of the vehicle application (thermal load, sulfur and other contaminations) over the aimed-for service life.

In addition, with an appropriate choice of components in such oxidation catalytic converters, nitrogen monoxide can be oxidized in accordance with reaction equation (E). As well as the importance of NO_2 to the SCR reaction (see Selective catalytic reduction) the emitted particulate matter can be reduced in accordance with the stoichiometry (F) with the created NO_2. Essentially, the hydrocarbons present in the particulates are oxidized using NO_2 in accordance with stoichiometry (D):

(E) $2\,NO + O_2 \rightarrow 2\,NO_2$,
(F) $C_x + 2x\,NO_2 \rightarrow x\,CO_2 + 2x\,NO$.

Without further exhaust-gas treatment measures it is not possible for a large number of vehicles to comply with strict particulate-emission limits (see for example EC Regulation 582/2011 [1]).

Filtration of particulates

To comply with the emission-control standard Euro 6c, a passenger car, irrespective of whether it is powered by a direct-injection gasoline engine or by a diesel engine, may not exceed the limit value for the particulate matter (PM) of 4.5 mg/km and the limit value for the particle number (PN) of $6 \cdot 10^{11}$ particles/km.

Particulate filter

Many vehicles need a particulate filter to achieve these limit values (Figure 1). These filters are for the most part ceramic honeycomb structures in which all the channels are alternately sealed at on end. Unlike other catalytic-converter honeycombs, it is therefore not possible to look through particulate filters. The exhaust gas must as it passes the honeycomb channels flow through the porous wall material (Figure 2). The pore structure of the ceramics is designed to accommodate the size of the particles. The most common particulate filters consist of silicon carbide (SiC), aluminum titanate (Al_2TiO_5), or cordierite ($Mg_2Al_3[AlSiO_5O_{18}]$). There are also special forms made of sintered metal.

Some of the residues of engine oil, fuel, combustion air, and engine attrition that collect in the particulate filter cannot be removed by combustion and, if suitably sized, remain as ash in the filter. On the other hand, soot and the organic compounds that adhere to it are combustible.

NO_2 and O_2 work as oxidizing agents. Oxidation with NO_2 works at exhaust-gas temperatures from around 200 to 250 °C. From around 450 °C NO prevails in the thermodynamic equilibrium of the nitrous fumes. Then there is little NO_2 in the exhaust gas.

Catalytically oxidizing coatings of particulate filters do not differ in their type and function fundamentally from other oxidation catalytic converters. They serve to set a favorable NO-NO_2 ratio, oxidize imperfectly combusted fuel, and catalyze the CO that is produced during soot burn-off.

Regeneration of the particulate filter

Soot is oxidized by oxygen only at temperatures above approximately 550 to 600 °C. Such temperatures exist in many diesel engines only with very high power outputs and hardly ever arise during the normal operation of some vehicles. Measures must therefore be taken to raise the exhaust-gas temperature. Combustion control of the engine, for example, is changed for this purpose.

Thermal measures

Important engine measures to increase exhaust-gas temperature are advanced burnoff or additive secondary injection, retarded main injection, retarded secondary injection, and intake-air throttling.

Figure 1: Diesel particulate filter
1 Housing,
2 Extruded honeycomb ceramic,
3 Ceramic plug.

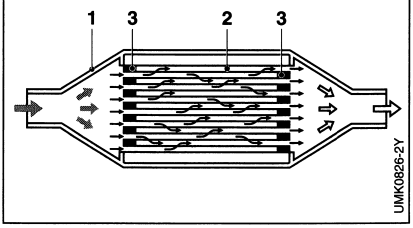
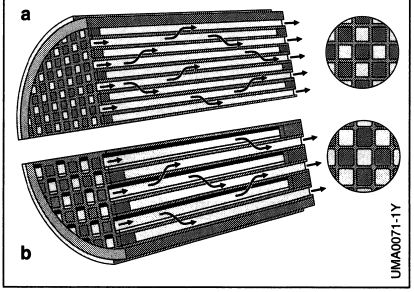


Figure 2: Designs of ceramic particulate filter
a) Square channel cross-section,
b) Octosquare design.



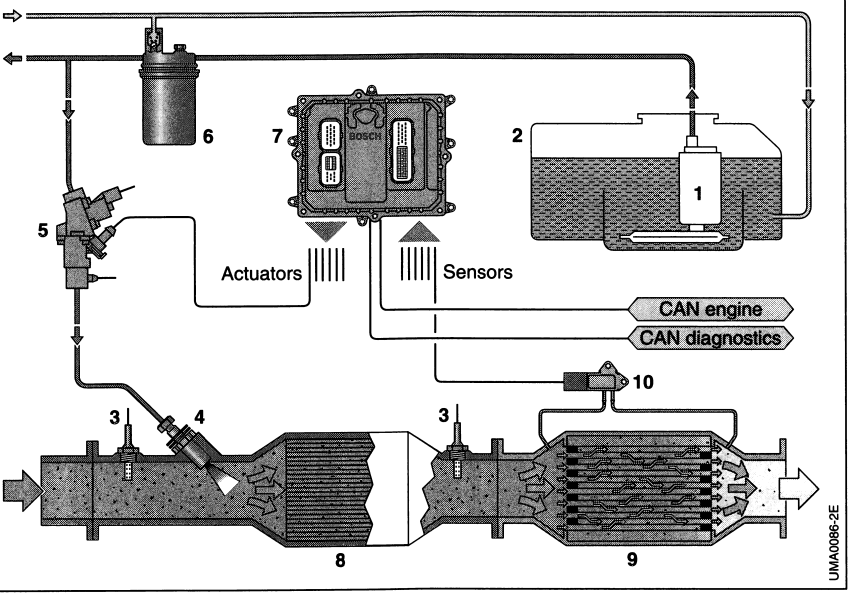
Depending on the engine operating point, one or several of these measures are triggered.

HCI system
As an alternative to injecting additional fuel into the engine cylinders, fuel can also be metered directly into the exhaust-gas system. The HCI system (Hydrocarbon Injection) serves this purpose. Here, the diesel fuel is for the most part injected ahead of an oxidizing catalytic converter and vaporized (Figure 3). The fuel vapors oxidize in the catalytic converter, resulting in an increase in the temperature of the exhaust-gas flow. The control algorithms of the HCI system are associated with soot-mass determination.

Additive system
In addition to the thermal measures to raise the exhaust-gas temperature, the ignition temperature of soot can also be reduced. This is done by catalytically active components which are finely distributed into the soot. These components are supplied as a fuel additive (Figure 4). Organometallic compounds are used, for example ferrocene. The finely distributed metal oxides formed from them reduce the ignition temperature of soot by up to 150 K. One drawback is the ash that is created from this fuel additive.

Saturation detection
Important factors in the design of particulate filters and their regeneration strategy are the amount of ash to be expected, the maximum soot load until regeneration is initiated, and the permissible exhaust-gas back pressure.

Figure 3: HCI system (Hydrocarbon Injection)
1 Fuel pump, 2 Fuel tank, 3 Temperature sensor, 4 HC metering module, 5 HC metering unit, 6 Fuel filter, 7 Engine ECU, 8 Diesel oxidation catalytic converter, 9 Diesel particulate filter, 10 Differential-pressure sensor.



Often two processes are used in parallel for saturation detection. The particulate filter's flow resistance is calculated from the pressure drop of the exhaust-gas flow detected with a differential-pressure sensor over the particulate filter. It can be related to the amount of soot stored in the particulate filter. In addition, a model is used to calculate the soot mass stored in the particulate filter. To this end, the soot-mass flow of the engine is integrated, where different corrections can be taken into consideration, e.g. for the soot oxidation of the particulates by NO_2.

Catalytic reduction of nitrous fumes

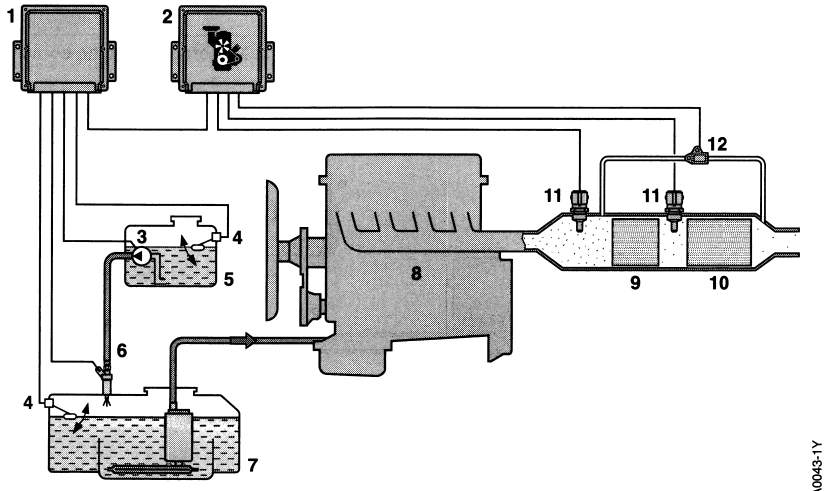
Requirements

Catalytic methods for reducing the NO_x emissions to comply with statutory requirements work particularly when vehicle design, engine application, and the components of the exhaust-gas treatment system are matched to each other. The following points must be taken into consideration here:

- The vehicle offers enough spacer to accommodate the necessary exhaust-gas treatment system. Special requirements with regard to the positioning of individual components in the vehicle must be observed (e.g. positioning near the engine).
- Exhaust-gas temperature, exhaust-gas mass flow, and exhaust-gas composition (NO_x, λ etc.) must be selectively and quickly regulated.
- The engine or a component of the exhaust system makes it possible to provide the amount of reducing agent

Figure 4: Particulate filter with additive system

1 Additive control unit, 2 Engine control unit, 3 Additive pump, 4 Fill level sensor, 5 Additive tank, 6 Additive injection nozzle, 7 Fuel tank, 8 Diesel engine, 9 Oxidation catalytic converter, 10 Particulate filter, 11 Temperature sensor, 12 Differential pressure sensor.



needed to reduce the untreated NO_x emissions.

The choice of a method of reducing NO_x suitable for a concrete vehicle application is therefore always a consideration of the complete system. The two NO_x storage catalytic converters and NO_x reduction catalytic converters described below can also be used together, for example to achieve high conversion rates.

NO_x storage catalytic converter

The reduction of NO_x from the exhaust gas of diesel engines works under suitable conditions with an NO_x storage catalytic converter (NSC, NO_x Storage Catalyst; alternatively called LNT, Lean NO_x Trap) very similar to the three-way catalytic converter (TWC, Three-Way Catalyst) (Figure 5).

Engine-based λ control constitutes an alternation of oxygen-richer and low-oxygen exhaust-gas sequences for this purpose. The oxygen-enriched phase often lasts several minutes. An accordingly large amount of storage components is required to bind the NO_x created during this period. The NO_x previously bound to the storage components is reduced in the subsequent fuel-enriched sequence. This sequence, known as NSC regeneration, usually takes a few seconds.

The storage components desorb NO_x at high temperatures also without the exhaust gas being enriched.

The NSC method works in particularly at exhaust-gas temperatures of between around 180 °C and 400 °C. The NO_x conversion of the NSC method deteriorates at significantly higher temperatures.

Contamination by sulfur reduces the efficiency of the noble metals and the NO_x storage components. A storage catalytic converter must therefore be regularly desulfurated. High temperatures are required for this purpose. When it comes to the design of NO_x storage catalytic converters, particular attention must be given to their long-term behavior depending on the relevant engine application with the details of desulfuration.

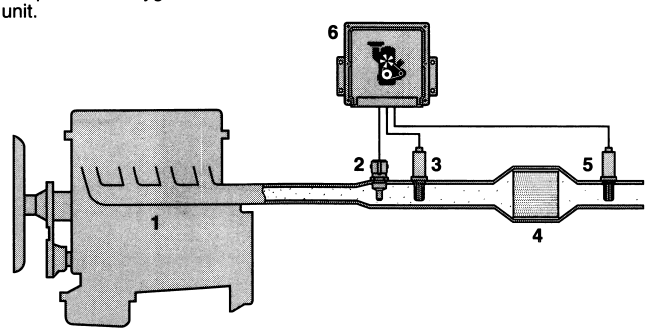
Selective catalytic reduction

SCR method

The SCR method (Selective Catalytic Reduction) is based on the principle of adding a reducing agent to the exhaust gas (Figure 6). Diesel fuel is basically also suitable here. However, it has hitherto not been possible to develop a suitable SCR method based on diesel fuel as the reducing agent that provides for compliance with strict emission limits.

Figure 5: Schematic diagram of an exhaust-gas system with NO_x storage catalytic converter

1 Diesel engine, 2 Temperature sensor,
3 Wide-band λ sensor (LSU),
4 NO_x storage catalytic converter,
5 NO_x sensor or two-position λ oxygen sensor,
6 Engine control unit.



Gaseous ammonia (NH_3) is used in stationary combustion systems, e.g. in power stations, in the SCR method. NH_3 gas is avoided in motor vehicles for safety reasons. Instead, in motor vehicles for the most part an aqueous solution of urea ($(NH_2)_2CO$) is metered into the exhaust-system branch in the direction of flow ahead of the SCR catalytic converter (Figure 6). Such aqueous solutions are called for example AdBlue, DEF (Diesel Exhaust Fluid), or AUS 32 (Aqueous Urea Solution) (see urea/water solution).

Aside from special forms, SCR catalytic converters frequently consist of catalytically active coatings on ceramic honeycomb structures. Special ceramic particulate filters are also suitable for such coatings. Vanadium oxides and iron and copper zeolites are suitable for use as the catalytically active material.

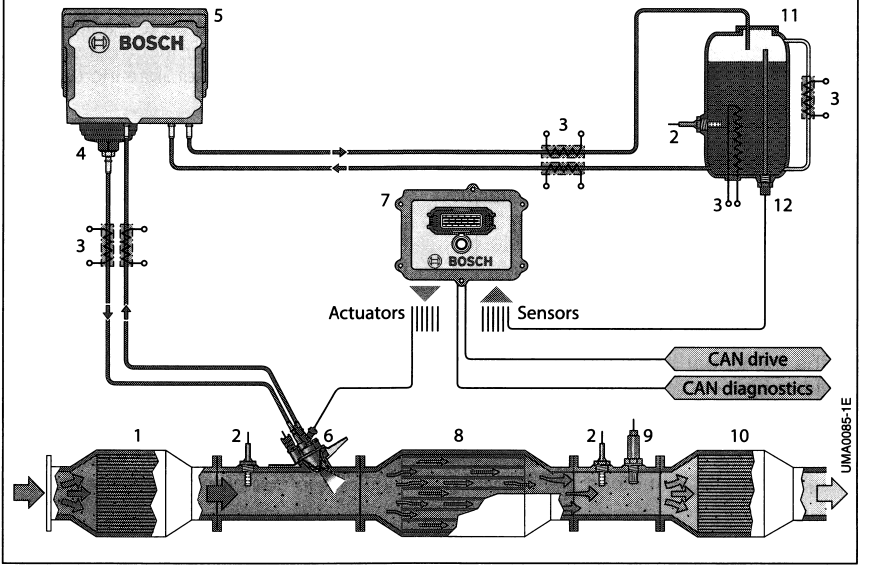
Reaction equations

Ammonia reacts on the SCR catalytic converter with NO_x according to the following equations:

(G) $4\ NO + 4\ NH_3 + O_2 \rightarrow 4\ N_2 + 6\ H_2O$,
(H) $NO + NO_2 + 2\,NH_3 + H_2O + O_2 \rightarrow 2\,NH_4NO_3$,
(I) $NH_4NO_3 \rightarrow N_2O + 2\ H_2O$,
(J) $2\ NH_4NO_3 \rightarrow 2\ N_2 + 4\ H_2O + O_2$,
(K) $NO + NO_2 + 2\ NH_3 \rightarrow 2\ N_2 + 3\ H_2O$,
(L) $6\ NO_2 + 8\ NH_3 \rightarrow 7\ N_2 + 12\ H_2O$.

The reaction according to equation (G) is comparatively slow. It predominates if no or only little NO_2 is available to react with NH_3. If the energy of activation is not achieved to reduce the nitrous fumes, ammonium nitrate (NH_4NO_3) is created according to equation (H) in a secondary reaction. NH_4NO_3 decomposes at temperatures from approximately 170 °C according to equation (I). If it is heated very quickly, it decomposes ex-

Figure 6: Exhaust-gas system with catalytic reduction of nitrogen oxides (SCR)
1 Diesel oxidation catalytic converter, 2 Temperature sensor, 3 Heater, 4 Filter, 5 Delivery module, 6 Metering module, 7 Metering control unit, 8 SCR catalytic converter, 9 NO_x sensor, 10 Slip catalytic converter, 11 Urea-solution tank, 12 Urea-solution fill-level sensor.



plosively according to equation (J). If the exhaust-gas flow contains equal proportions of NO and NO_2, the faster reaction predominates according to equation (K) if its energy of activation is reached. Where there is a higher proportion of NO_2 in the exhaust gas, the reaction can according to equation (L) contribute significantly to the overall conversion.

Choice of SCR catalytic converter
The engine emissions and the NO_x conversion needed to achieve the emission limit and the exhaust-gas temperature in the relevant operating states are decisive factors in the choice of a suitable SCR catalytic converter. However, the sulfur content of the fuel used is also important. When it comes to the design of the SCR catalytic converter, attention must also be given in particular to its long-term stability depending on the concrete vehicle application, e.g. also with the details of desulfuration.

Metering of the reducing agent
SCR catalytic converters do have a certain NH_3 storage capacity, but essentially a quantity of urea/water solution equivalent to the instantaneous NO_x mass flow must be metered as precisely as possible and dispensed into the exhaust-system branch. If more reducing agent is dispensed than is converted in reduction with NO_x, it may result in unwanted NH_3 slip. This excess NH_3 can be oxidized by an additional oxidation catalytic converter at the end of the SCR catalytic converter. In the most favorable case, nitrogen and water are created at this NH_3 slip catalytic converter. Nitrogen oxides are created in secondary reactions in unfavorable circumstances.

Before entering each catalytic converter, the exhaust-gas mass flow must be distributed as uniformly as possible to the individual honeycomb channels. Uniform distribution of the fluids requires particular attention in the case of two-substance mixtures, i.e. if a second substance is added to the exhaust-gas flow, as is the case with the SCR method.

To facilitate the SCR reactions, ammonia must first be created from the aqueous urea solution. This occurs in a chronological sequence of different steps:
- determination of the metering quantity,
- delivery of the solution,
- dispensing of the metered volume,
- atomization and vaporization of the solution,
- decomposition of urea, and
- distribution of ammonia through the catalytic-converter cross-section.

The metering quantity is determined in a separate ECU of the exhaust-gas treatment systems (optionally integrated in the engine control unit) on the basis of stored data and data determined by exhaust-gas sensors (temperature sensor, NO_x sensor, NH_3 sensor, etc.). In the process the current NH_3 storage load of the catalytic converter is also calculated and taken into account as a rule. The data required for this calculation are determined in advance for the catalytic converter used in each case and stored as a program map in the ECU. To meter the required quantity as precisely as possible, it is necessary to bear in mind that ISO 22241-1 [2] defines the urea content of these solutions at 31.8 to 33.2 % by mass. Significant deviations in the urea-solution concentration can also arise as a result of their aging in the vehicle or incorrect storage. Quality sensors are therefore increasingly being used.

Delivery and metering modules
A delivery module as a rule consists of a liquid pump, e.g. a diaphragm pump. It provides the solution from the urea tank for metering. Since commercially available urea solutions freeze below −11 °C, all the components that are in contact with the solution must be able to withstand ice pressure. Some components have heating elements to melt the ice where necessary. The metered volume is dispensed via metering modules. These are primarily liquid nozzles, or more rarely air/fluid two-substance nozzles. The metering modules in contact with the hot exhaust

gas must be cooled to prevent the solution from vaporizing prematurely.

Vaporization of the solution is preceded by the distribution of the droplets sprayed into the exhaust-gas flow. The spray characteristic, i.e. the droplet size and its impulse, must be closely matched with temperature and flow conditions in the affected part of the exhaust-system branch. As a rule some of the droplets vaporize on hot surfaces of the exhaust system, e.g. on pipe walls or on gas mixers.

Release of ammonia

Vaporization of the droplets is followed by the release of ammonia due to the decomposition of urea. In the first thermolytic decomposition stage of urea, initially isocyanic acid (HNCO) is created as well as NH_3 under the conditions prevailing in the exhaust-system branch according to reaction equation (M):

(M) $(NH_2)_2CO \rightarrow NH_3 + HNCO$.

This decomposition of urea occurs above its melting temperature from approximately 133 °C. Under suitable conditions the isocyanic acid created in the first stage is subjected in a second stage to hydrolysis with the creation of ammonia and carbon dioxide according to equation (N):

(N) $HNCO + H_2O \rightarrow NH_3 + CO_2$.

This comparatively slow hydrolysis reaction is significant in the gaseous phase merely in the case of a long retention time of very small droplets, which vaporize already in the gaseous phase at accordingly high temperatures. Hydrolysis of isocyanic acid therefore occurs for the most part on surfaces. As well as the most uniform distribution as possible of the created ammonia into the exhaust-gas flow and its distribution as perfectly as possible through the catalytic-converter cross-section, it must also be borne in mind when designing that part of the exhaust-system branch into which the agent is metered that solid by-products may be created when urea decomposes. Surface temperatures of well over 300 °C are needed to eliminate their deposits.

On-Board Diagnostics (OBD)

OBD describes self-diagnostic systems for testing essential vehicle functions, including exhaust-gas treatment. A multitude of sensors and comprehensive software supply most of the information for these diagnoses. OBD has been subject to continuous further development over the past (more than) 20 years. Today OBD systems influence the vehicle's service-inspection program and the driver and the workshop are informed of malfunctions and recommended actions. The OBD requirements applicable in each case must be taken into account in the design of the exhaust-gas treatment system.

References

[1] Commission Regulation (EU) No. 582/2011 of 25 May 2011 implementing and amending Regulation (EC) No. 595/2009 of the European Parliament and of the Council with respect to emissions from commercial vehicles (Euro VI) and amending Annexes I and III to Directive 2007/46/EC of the European Parliament and of the Council (Text with EEA relevance).

[2] ISO 22241-1: Diesel engines – NO_x reduction agent AUS 32 – Part 1: Quality requirements.

Conventional drivetrain

Drivetrain elements

The conventional drivetrain of a motor vehicle includes the internal-combustion engine, the manual transmission, the propeller shaft, the differential, and the drive shafts. The manual transmission contains a vibration damper, a power take-up element, a step-up gear unit with gear-shift elements, and if necessary a transfer case (Figure 1).

Further drivetrain arrangements, in which an electric machine operates in addition to the internal-combustion engine, are discussed elsewhere (see Drivetrain with hybrid drives).

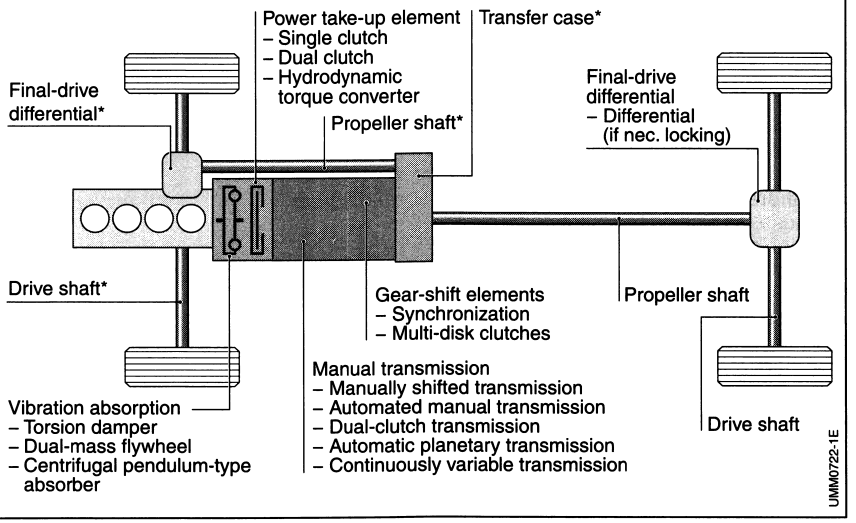
Function

The function of a drivetrain is to adapt the power outputs and torques made available by the engine to the respective driving situation and to the driver's commands.

Each internal-combustion engine converts the energy of the fuels supplied to it into mechanical energy, operates between a minimum and a maximum rotational speed in the process, and delivers a limited torque in this range. If an internal-combustion engine were to power the drive wheels without the transmission ratio of manual transmission and differential, then for a passenger car the engine's idle speed would correspond to a road speed of approximately 80 km/h and the maximum engine speed to a road speed of approximately 500 km/h.

Unlike steam engines or electric machines, the internal-combustion engine cannot generate torque when it is at a standstill. Moreover, fuel consumption is very much dependent on the requested torque at the applied engine speed. Consequently, this gives rise to the following functions on the part of a drivetrain:

Figure 1: Drivetrain elements for the standard drive
Components marked * are optionally fitted.



- converting the engine torque and the engine speed to match the respective driving situation,
- ensuring idle operation of the internal-combustion engine when the vehicle is at a standstill,
- facilitating driving off from a standstill,
- isolating the drivetrain from the rotational irregularities of the internal-combustion engine,
- reversing the direction of rotation of the drive wheels where necessary (reverse gear).

The requirements that a motor vehicle is to fulfill with regard to maximum climbing capability and acceleration on the one hand and top speed on the other hand define the necessary tractive force on the wheels and hence the demand on the transmission with regard to its spacing between the lowest and highest gears.

The tractive-force hyperbola (Figure 2) shows as a function of the driving speed v and the drivetrain efficiency η_A the maximum available tractive force F_Z for the event that the internal-combustion engine could output its maximum power P_{max} over its entire speed range:

$$F_Z = \frac{P_{max}}{v}\,\eta_A.$$

In the lower speed range the tractive force is limited not by the tractive-force hyperbola but by the adhesion limit F_{Zmax} between tire and road which is obtained from the product of the normal force F_N of the driven vehicle wheels and the coefficient of friction μ of the roadway:

$$F_{Zmax} = F_N\,\mu.$$

The running resistance F_W of the vehicle without acceleration is made up of rolling resistance F_R, aerodynamic drag F_L, and climbing resistance F_S. The factors that influence the rolling resistance are the vehicle mass m, the gravitational acceleration g, the coefficient of friction μ, and the gradient angle α of the roadway. The aerodynamic drag contains the drag

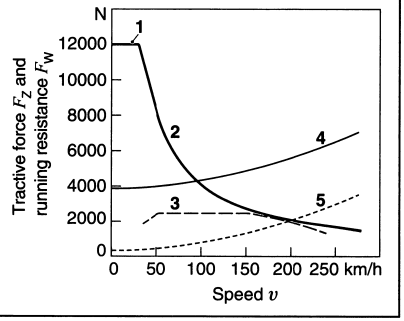
Table 1: Quantities and units

Quantity		Unit
c_W	Drag coefficient	–
g	Gravitational acceleration	m/s²
i_n	Transmission ratio of a gear	–
m	Vehicle mass	kg
n	Revolutions per minute	rpm
v	Driving speed	m/s
A	Frontal area of vehicle	m²
F_Z	Tractive force at wheels	N
F_N	Normal force of vehicle	N
F_W	Running resistance	N
F_R	Rolling resistance	N
F_L	Aerodynamic drag	N
F_S	Climbing resistance	N
F_{Syncr}	Axial force in synchronization	N
M	Torque	N m
P	Engine power	kW
α	Gradient angle of roadway	degree
γ	Torque-conversion factor	–
φ	Gear step	–
φ_G	Transmission spacing	–
η	Efficiency	–
μ	Coefficient of friction of roadway	–
ρ	Density of the air	kg/m³

Subscripts:
min minimum
max maximum
A drivetrain
G transmission

Figure 2: Tractive force as a function of speed for an internal-combustion engine without manual transmission

1 Adhesion limit F_{Zmax},
2 Tractive-force hyperbola F_Z,
3 Tractive force F_Z for engine without manual transmission,
4 Characteristic running-resistance curve at $\alpha = 10°$,
5 Characteristic running-resistance curve at $\alpha = 0°$.



UMM0723E

coefficient c_W, the frontal area of the vehicle A, the density ρ of the air, and the driving speed v. The factors that influence the climbing resistance are the vehicle mass m, the gravitational acceleration g, and the gradient angle α of the roadway:

$$F_W = F_R + F_L + F_S,$$

$$F_W = mg\mu\cos\alpha + 0.5\,c_W A\rho v^2 + mg\sin\alpha.$$

In Figure 2 it becomes clear that an internal-combustion engine without a manual transmission would only operate in a very limited driving range. Here the tractive force F_Z of the engine only contains the transmission ratio of the differential so that the effect of torque conversion in the manual transmission can be clearly explained in the following.

Manual transmission

A manual transmission effects an approximation of the maximum engine torque – which the internal-combustion engine develops between its minimum and maximum speeds – to the tractive-force hyperbola. Highest and lowest gear ratios i define the spacing φ_G of the transmission:

$$\varphi_G = \frac{i_{min}}{i_{max}}.$$

Both geometrical and progressive gear stepping can be implemented with the number of gears and the individual transmission ratios. The gear step φ between the gears is defined as the ratio of transmission i of two neighboring gears:

$$\varphi = \frac{i_{n-1}}{i_n}.$$

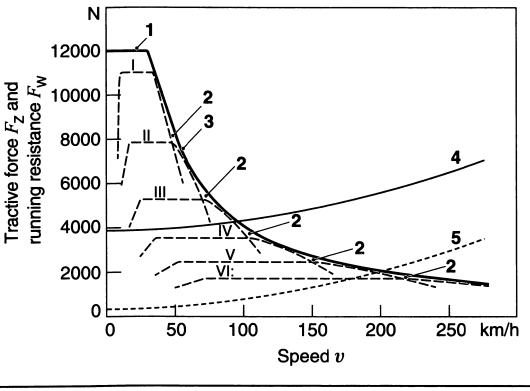
In the case of geometrical gear stepping the gear step is always the same. This is used in truck transmissions with high numbers of gears. A truck requires fine stepping in the entire driving range since its engine power is relatively low with regard to its high mass.

Progressive gear stepping is used in passenger-car transmissions. This means that the gear step becomes increasingly smaller to the higher gears. Consequently, the non-utilizable driving ranges (see Figure 3) are kept small in the car's main driving range although the number of gears is much smaller than in a truck. To achieve good driveability the tractive-force curves of the individual gears should intersect in each speed range (see Figure 3).

The point of intersection between the characteristic running-resistance curve at a 0° gradient angle and the tractive-force hyperbola defines the vehicle's top speed v_{max}. The ratio of the gear in which the top speed is to be reached should be selected in such

Figure 3: Tractive force as a function of speed for an engine with six gear steps (v_{max} in fifth gear, plus an overdrive gear)

1 Adhesion limit F_{Zmax},
2 Non-utilizable driving range,
3 Tractive-force hyperbola F_Z,
4 Characteristic running-resistance curve at α = 10°,
5 Characteristic running-resistance curve at α = 0°.
I...VI Tractive-force curves for gears 1 through 6.



a way that the engine outputs its maximum power at this point (fifth gear in Figure 3).

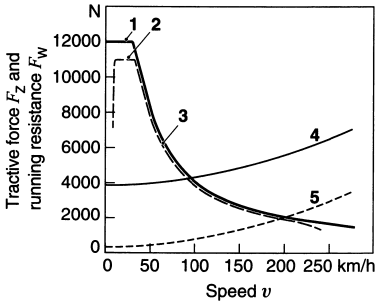
To reduce fuel consumption there are one or more gears with an overdrive ratio (sixth gear in Figure 3). In this gear the top speed cannot be reached (in Figure 3 a maximum speed of approximately 180 km/h on a 0° gradient is possible), but reducing the engine speed lowers the fuel consumption in a wide road-speed range.

Continuously variable transmission

A continuously variable transmission can make any number of transmission ratios available. In this way, the maximum engine tractive force is introduced to the tractive-force hyperbola (Figure 4). On account of their design, continuously variable transmissions have a limited transmission spacing and a lower transmission efficiency. In Figure 4 the highest ratio corresponds roughly to the ratio of 5th gear of the multi-step transmission in Figure 3. Consequently, above all in the upper speed ranges, there is no possibility of reducing the engine speed and with it fuel consumption.

Figure 4: Tractive force as a function of speed for an engine with continuously variable transmission

1 Adhesion limit F_{Zmax},
2 Tractive force F_Z for continuously variable transmission,
3 Tractive-force hyperbola F_Z,
4 Characteristic running-resistance curve at $\alpha = 10°$,
5 Characteristic running-resistance curve at $\alpha = 0°$.



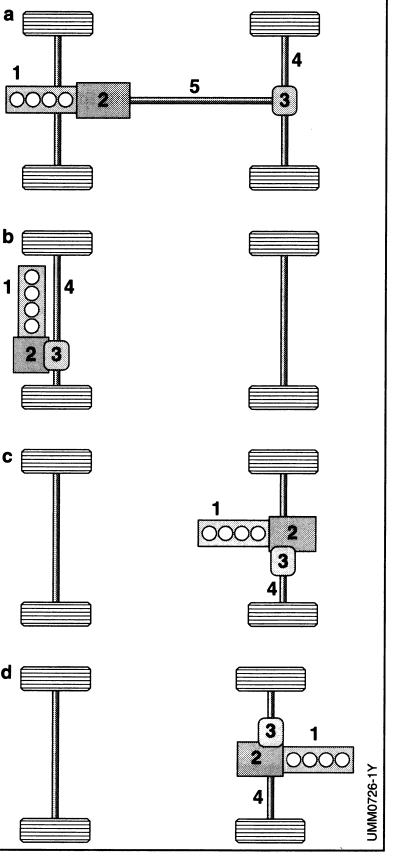
UMM0725E

Arrangement of engine

The arrangement of the engine and the manual transmission depends very much on the vehicle category. Until around 1930 virtually all vehicles had the engine at the front and the driven wheels at the rear. The designation standard drive for this config-

Figure 5: Arrangement of engine

a) Standard drive (longitudinal arrangement of engine),
b) Front-wheel drive (transversal arrangement of engine),
c) Mid engine,
d) Rear engine.
1 Engine,
2 Manual transmission,
3 Final-drive differential,
4 Drive shaft,
5 Propeller shaft.



UMM0726-1Y

uration derives from this era (Figure 5a). Since then, this type of drivetrain has featured almost exclusively in the upper mid-size and luxury categories. The four-wheel drive of many off-road vehicles is a variant of the standard drive with a transfer case situated in the direction of power flow after the manual transmission which transfers the transmission output torque to the front and rear axles (see Figure 1).

The front-wheel drive has in the meantime gained acceptance in vehicles ranging from the subcompact-size to the lower mid-size category and now features in more than 75 % of worldwide vehicle production. Here the engine and the manual transmission are arranged transversally to the direction of travel in the engine compartment (Figure 5b). The final-drive differential is integrated in the manual transmission, hence there is no propeller shaft. There are also versions of the front-wheel drive with four-wheel drive.

The mid engine (Figure 5c) and the rear engine (Figure 5d) with rear-wheel drive are now installed almost exclusively in sports vehicles.

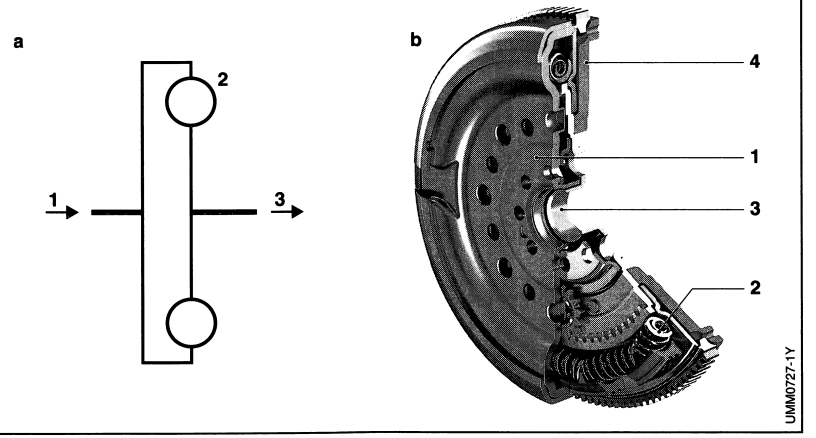
Vibration isolation

A significant requirement imposed on the drivetrain is comfort. The requirements imposed on vibration isolation are increasing continuously since optimizations in the combustion process of engines, smaller engine displacements with turbocharging, lower engine speeds in all driving ranges, and reduced numbers of cylinders cause a higher rotational irregularity of the combustion-engine torque.

The purpose of vibration isolation is to reduce the torsional vibrations that occur as a result of the design of the internal-combustion engine. Vibration isolation can be effected by slip at the power take-up element, by damping the irregularity, by temporarily storing energy in a spring, by absorption, or by a combination of several methods.

In the case of vibration isolation by slip at the power take-up element, the mechanical energy of the vibration by friction is converted into heat. This energy is thus lost to the actual drive. Therefore this vibration-isolation option is now only used sporadically in critical vibration situations (e.g. low rotational speeds).

Figure 6: Torsion damper and dual-mass flywheel
a) Schematic diagram,
b) Structural design.
1 From driving engine,
2 Coil springs,
3 To power take-up element,
4 Flywheel mass.



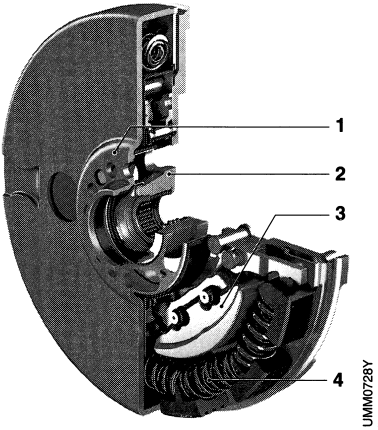
Torsion damper and dual-mass flywheel

The torsion damper consists of coil springs arranged in a circle between the engine and the power take-up element (Figure 6), which during the vibration temporarily store energy in the compression of the springs (by a limited torsion between crankshaft and transmission inlet) and output this energy again to the transmission inlet. It is possible by varying the spring stiffness and diameter to adapt the torsion damper to the respective combination of engine and drivetrain.

In the case of the dual-mass flywheel the mass of the engine flywheel is apportioned and the torsion damper is positioned between the two flywheel masses. The moment of inertia on the transmission side is increased by apportioning to the primary flywheel mass on the engine side and to the secondary flywheel mass on the transmission side. Low spring stiffness can bring about a specifically "soft" harmonization of the spring units which connect the two flywheel masses to each other. As a result, the resonant frequency of the dual-mass flywheel is shifted significantly below the idle speed of the engine and the exciting engine orders. A natural oscillation of the dual-mass flywheel is avoided in this way.

Figure 7: Centrifugal pendulum-type absorber and torsion damper

1 From driving engine,
2 To power take-up element,
3 Centrifugal pendulum-type absorber mass,
4 Coil spring.



Centrifugal pendulum-type absorber

In the case of the centrifugal pendulum-type absorber the rotational irregularity is balanced by means of a counterforce by vibrating masses additionally mounted in the torsion damper (Figure 7). Thanks to suitable lever conditions the vibrating masses can be selected at a proportionally small level compared with the mass of the internal-combustion engine to be isolated. The centrifugal pendulum-type absorber is speed-adaptive since the centrifugal acceleration on the vibrating mass increases as the rotational speed increases.

Power take-up elements

The internal-combustion engine exhibits a minimum speed. To drive off from a vehicle standstill, the rotational-speed gap between the lowest engine operating speed and the stationary transmission input shaft must be closed by a speed converter [1]. The functions of this power take-up element are:
- interrupting the power flow between the engine and the step-up gear unit,
- soft and smooth starting,
- transmitting the engine torque at different engine speeds to the transmission.

Single clutch

Single-disk clutch

The single-disk clutch installed in most manual transmissions consists of a clutch disk fitted with a clutch lining (Figure 1) which is pressed by a spring-pretensioned pressure plate against the engine flywheel (Figure 2). Without an external operating force the spring force keeps the clutch closed. To open the clutch, it is necessary to depress the clutch pedal.

The clutch pedal has a mechanical or hydraulic connection to the release fork on the clutch. The release fork counteracts the spring pre-tension of the pressure plate via the throwout bearing. In this way, as the pressure on the clutch pedal increases, the transmitted torque decreases until the clutch is fully opened. When during driving off the force is slowly withdrawn via the clutch pedal, the spring increasingly presses the pressure plate – via the spring – against the engine flywheel. The transmitted torque increases and the vehicle drives off. When speed synchronization between the engine and the transmission has taken place, the drive-off process is completed and the clutch is fully engaged.

Figure 1: Clutch disk with clutch lining
1 Clutch lining,
2 Clutch disk,
3 To transmission.

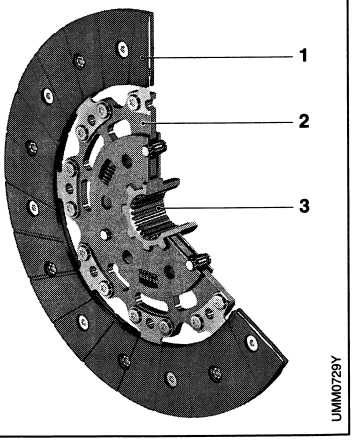
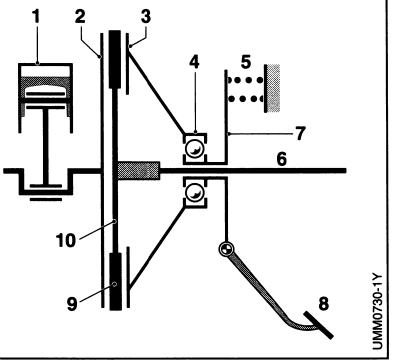


Figure 2: Operation of a single clutch
1 Internal-combustion engine,
2 Engine flywheel, 3 Pressure plate,
4 Throwout bearing, 5 Spring,
6 Transmission input shaft,
7 Release fork, 8 Clutch pedal,
9 Clutch lining, 10 Clutch disk.



Double-disk clutch

In rare cases there are versions for higher torques with two clutch disks (Figure 3). This version would be chosen if the single-disk clutch were too large in diameter. Both clutch disks act on only one transmission input shaft. For this reason, the double-disk clutch too belongs to the category of single clutches.

Figure 3: Double-disk clutch
1 From driving engine, 2 To transmission,
3 Clutch disk 1 with lining,
4 Clutch disk 2 with lining.

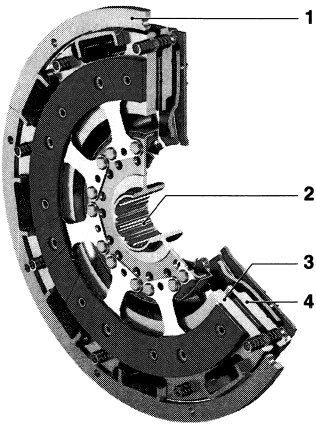
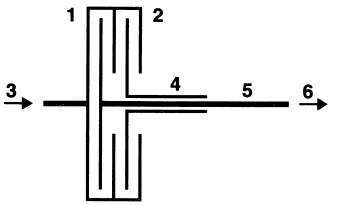


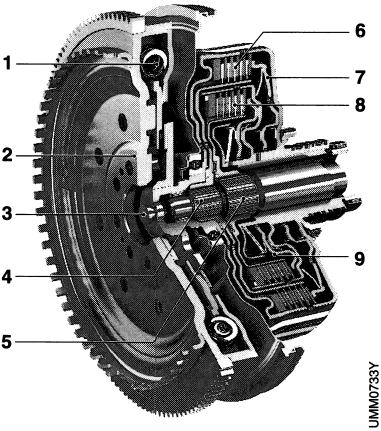
Figure 4: Schematic diagram of a dual clutch
1 Clutch 1, 2 Clutch 2,
3 From vibration isolation, 4 Hollow shaft,
5 Solid shaft, 6 To transmission.



Dual clutch

The dual clutch (Figure 4 and Figure 5) is characterized by the fact that two independently controllable clutches transmit the engine power to two transmission input shafts (a solid shaft and a hollow shaft). In most cases the dual clutch is combined with a dual-clutch transmission (see Dual-clutch transmission). The even gears are assigned to one clutch while the odd gears are assigned to the other. In this way, one clutch can transmit torque in a gear while the second clutch is opened for pre-selection of the next gear. During the gear shift the torque is transmitted without an interruption of tractive force from the first to the second clutch. The other clutch reduces its torque transmission at the extent to which the clutch to be switched increases its transmitted torque. The gear shift takes place in slipping mode while the dual clutch functions as a power-shift element.

Figure 5: Structural design of a dual clutch
1 Torsion damper, 2 From driving engine,
3 From vibration isolation,
4 Solid shaft (clutch 1),
5 Hollow shaft (clutch 2),
6 Clutch 1, 7 Operation clutch 1,
8 Clutch 2, 9 Operation clutch 2.



Dry and wet clutches

Single and dual clutches can take the form of both a dry clutch and an oil-cooled "wet" variant.

Dry clutch
Clutch linings in dry clutches are made of organic materials which are bonded by resins. The materials used are similar to those of brake pads/linings. The aim is to achieve a high coefficient of friction for the clutches so as to keep the operating force and the diameter of the clutch small.

Wet clutch
Multi-disk clutches with paper-like linings are used for wet clutches. These clutches are similar to the power-shift elements in automatic transmissions (see Gear-shift elements). The function of the paper-like lining is obtained by oil wetting. More heat can be dissipated via an oil flow through grooves in the clutch disks from the clutch than is the case with dry linings, where cooling is effected by air convection.

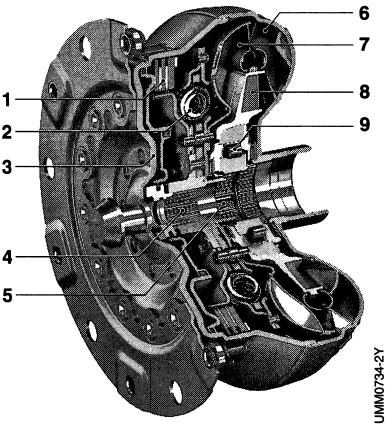
The oil-cooled multi-disk clutches (see Figure 5) are usually actuated not via throwout bearings but by co-rotating pistons which generate an axial force by means of hydraulic pressure. When a multi-disk clutch is compressed with increasing axial force, the torque that can be transmitted via the clutch increases.

Hydrodynamic torque converter

The hydrodynamic torque converter (Figure 6) comprises the following components:
- impeller (connected to the internal-combustion engine),
- turbine (connected to the transmission input shaft),
- stator (connected via a one-way clutch to the transmission housing),
- lockup clutch, and
- integrated torsion damper.

The torque converter is filled with oil. The kinetic energy introduced by the engine is converted in the impeller into hydraulic energy and applied in the turbine again in kinetic energy to the transmission input shaft. Because the conversion takes place using the hydrodynamic inertia of the oil, the hydrodynamic torque converter

Figure 6: Structural design of hydrodynamic torque converter
1 Lockup clutch,
2 Torsion damper,
3 From driving engine,
4 To transmission input shaft,
5 Connection to transmission housing,
6 Impeller,
7 Turbine,
8 Stator,
9 One-way clutch.



operates without wear. The torque at the impeller is called pump torque M_P, that at the turbine is called turbine torque M_T. The faster the impeller driven by the engine rotates at speed n_P, the more oil is delivered by the impeller from the inside close to the rotational axis radially outwards. This oil flow strikes vanes in the neighboring turbine. The deflection of the oil flow to these vanes induces a torque which acts on the transmission input shaft.

The oil flow is delivered in the turbine from the outside radially inwards, where it strikes the stator (Figure 7). The stator features vanes which have a curvature of 90° and are supported by means of a one-way clutch in one direction of rotation on the transmission housing. While the turbine (and hence the transmission inlet) is stationary or only rotates slowly at speed n_T, the stator vanes induce a retention of oil which results in a reaction-torque increase at the turbine. With the oil flow being supported by means of the stator with stator torque M_L on the transmission housing, it is possible that a higher torque is transmitted to the transmission input shaft than the internal-combustion engine delivers to the converter. This is referred to as a converter increase. The following torque equilibrium applies to the torque converter:

$$M_P + M_T + M_L = 0$$

With regard to the torque converter the pump torque is an input torque and hence the preceding sign is positive, while the turbine torque is an output torque and hence the preceding sign is negative.

As the speed of the turbine (i.e. the transmission input shaft) increases, the angle of impact of the oil flow on the stator vanes changes. The oil retention and hence the torque supported on the transmission

Figure 7: Hydrodynamic torque converter
a) Schematic diagram in cross-section,
b) Oil-flow course.
The turbine and the impeller rotate in such a way that the visible front moves downward. The stator, which would like to move in the conversion range in the opposite direction, is supported via a one-way clutch on the transmission housing.
The round arrows in the figure on the left symbolize the direction of the oil flow.
1 Turbine, 2 Stator, 3 Impeller, 4 Hollow shaft for stator support,
5 Housing (fixed), 6 Hollow impeller shaft for driving the transmission oil pump,
7 Converter cover connected to impeller, 8 Stator one-way clutch,
9 Turbine shaft (transmission inlet), 10 Turbine direction of rotation,
11 Oil flow direction, 12 Impeller direction of rotation.

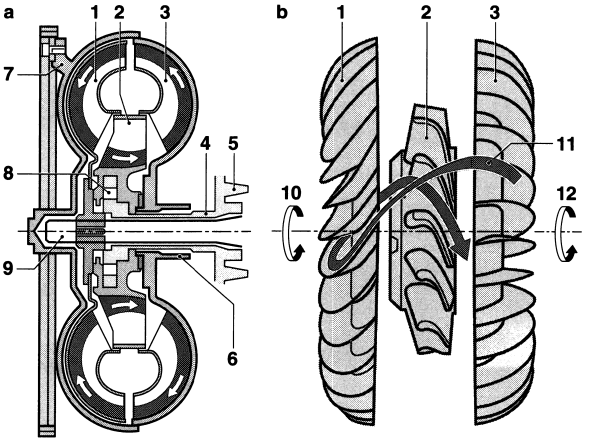
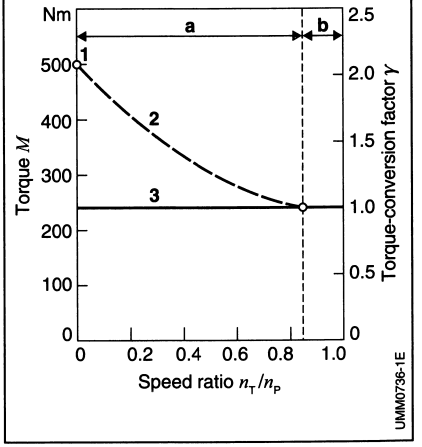


Figure 8: Torque-conversion factor in hydrodynamic torque converter
1 Stall point,
2 Torque M_T at turbine or torque-conversion factor γ,
3 Drive torque M_P at impeller.
a Conversion range,
b Clutch range.
n_T Turbine speed,
n_P Impeller speed.



housing are reduced. The converter increase decreases. The characteristic can be seen in Figure 8 in the torque-conversion factor γ:

$$\gamma = -\frac{M_T}{M_P}.$$

At the stall point when the turbine is stationary $1.7 < \gamma < 2.4$ usually applies.

As long as $\gamma > 1$ applies the converter is in the so-called conversion range. The range where $\gamma = 1$ is designated the clutch range since no further torque conversion takes place here.

When the torque-conversion factor γ has dropped to 1, the oil flow acts on the stator vanes against the previous torque direction. The one-way clutch, which constitutes the connection between the stator and the transmission housing, transmits no torque in this direction. The stator now also rotates without torque in the oil flow. The co-rotation of the stator prevents an unwanted deflection of the oil flow in the clutch range and consequently a deterioration of the converter efficiency.

To keep the losses in the converter at $n_T = n_P$ as low as possible, hydrodynamic transmission is locked up by a multi-disk clutch. To maintain the high vibration comfort of hydrodynamic isolation, one or more torsion dampers are incorporated in the lockup clutch's power-transmission path.

References

[1] H. Naunheimer; B. Bertsche; G. Lechner: Fahrzeuggetriebe – Grundlagen, Auswahl, Auslegung und Konstruktion. 2nd Ed., Springer-Verlag, 2007.

Transmission

Transmissions are a combination of step-up gear units and gear-shift elements. The vibration-damping and the power take-up elements are arranged between the engine and the transmission.

Figure 1: Involute of a circle
1 Base circle, 2 Involute,
d_b Base-circle diameter.

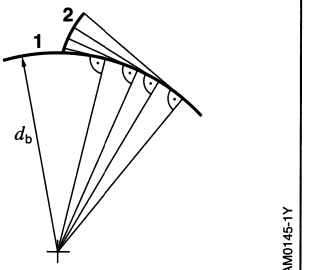
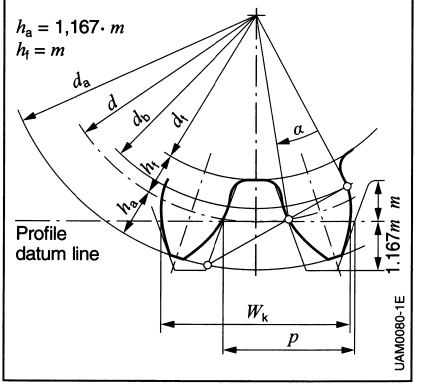


Figure 2: Parameters for a gear
d_a Outside diameter,
d Pitch diameter,
d_b Base-circle diameter,
d_f Root diameter,
h_a Addendum,
h_f Dedendum,
W_k Base tangent length,
p Pitch $p = \pi m$;
m Module,
α Pressure angle.



Step-up gear units

Different gear ratios are implemented in the manual transmission for the purpose of adapting the torque made available by the engine and the power to the respective driving situation and to the driver's requests. Each gear ratio is called a gear or speed.

Gears

The gear (or gear wheel) is a machine element which has uniformly distributed teeth around its circular circumference. Torque and rotary motion are transmitted by a positive engagement occurring in succession of the teeth of a gear with the teeth of other gears. The gear ratio during the transmission from one gear to another is determined by different numbers of teeth on the gears. Rotational speeds and torques are inversely proportionally converted with this gear ratio, with the power output apart from efficiency remaining the same. The advantages of a step-up gear train lie in its high power density through slip-free transmission.

The type of toothing most commonly used in step-up gear units is involute toothing. Here the cross-section of a tooth flank depicts a part of an involute of a circle (Figure 1). This ensures that the engaging teeth touch each other along a straight line of action. The aim of involute toothing is to achieve a high amount of rolling and an accordingly low amount of sliding during the transmission of torque in order to minimize wear and heat generation in the toothing. Figure 2 shows the parameters for a gear.

Straight and helical toothing

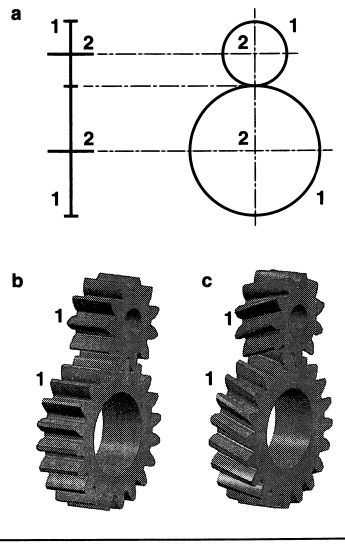
Helical gears are used for the most part in spur-gear drives. On these gears the teeth are arranged not parallel with the gear axes, but at an angle to them. When the teeth of two gears mesh, the teeth do not bear on their full face width from the outset, as is the case with straight spur gears. Instead, the bearing face width increases slowly as the gears continue to rotate until the teeth have contact over their full width. As the gears rotate out of the action zone the face width subjected to load also

decreases slowly. In this way, the kinetic impulses when meshing are much smaller than on straight gears. Accordingly, fewer oscillations are excited, which provides for quieter running of the gears.

Acoustically, it is possible to discern the difference between straight and helical toothing in older manually shifted transmissions. In such transmissions the reverse gear is straight-toothed and "whines" when reversing, while the forward gears are acoustically inconspicuous when driving forwards.

Helical toothing generates axial forces which must be absorbed by means of a suitable bearing arrangement. Moreover, they are more complex to produce than straight spur gears.

Figure 3: Spur-gear set
a) Schematic diagram,
b) Structural design of a straight spur-gear set,
c) Structural design of a helical spur-gear set.
1 Spur gear,
2 Axis.



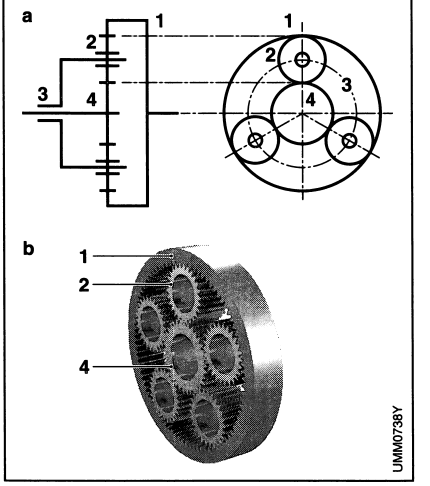
Spur-gear set
In spur-gear sets the gears are arranged on parallel axes (Figure 3). The gears can either be permanently coupled on shafts (fixed gear) or pivoted on them (movable gear).

Spur gears are easier to manufacture than internally toothed ring gears of a planetary-gear set. Moreover, the axial offset between input and output suggests itself for front-transverse drivetrains.

Spur-gear sets are used in manually shifted transmissions, automated manual transmissions, and dual-clutch transmissions.

Planetary-gear set
The planetary-gear set consists of a sun gear, an internally toothed ring gear, and several planetary gears. The planetary gears are interconnected by a carrier and mounted on this (Figure 4). The planetary-gear set, in contrast to the spur-gear set, has not only fixed axes. The axes of

Figure 4: Planetary-gear set
a) Schematic diagram,
b) Structural design.
1 Ring gear,
2 Planetary gear,
3 Carrier,
4 Sun gear.



the planetary gears can circulate in an orbit around the axes of the sun and ring gears, like the planets around the sun in our solar system.

Basically, in each case one of the three elements (sun gear, ring gear, carrier) in the planetary-gear set can be used as the drive, while a second element is permanently joined to the transmission housing in order to provide output via a third element. With three drive options and three output options (with the third element stationary in each case) there are six different gear ratios which a planetary-gear set can offer when connected appropriately. Furthermore, by joining the three elements to each other, it is possible to achieve a circulation of the planetary-gear set as a block with the gear ratio one as the seventh ratio. However, the ratios of a planetary-gear set are not so staged that a planetary-gear set alone would be usable for the desired ratios of a passenger-car transmission.

A planetary-gear set can also be used as a summation gearbox or transfer case. In this case, none of the planetary-gear set elements are joined to the housing. When it is used as a summation gearbox, two elements of the planetary-gear set are driven and output via the third element. When it is used as a transfer case, one of the planetary-gear set elements operates as the drive while the other two operate as the output.

Thanks to the symmetrical design and the distribution of torque to several planetary gears, the power density in the planetary-gear set is particularly high and accordingly the size in relation to the transmitted torque is small. The coaxial position of the drive and output shafts make this construction type attractive for standard drivetrains. Several gear sets combined via shift elements are predominantly used in automatic planetary transmissions.

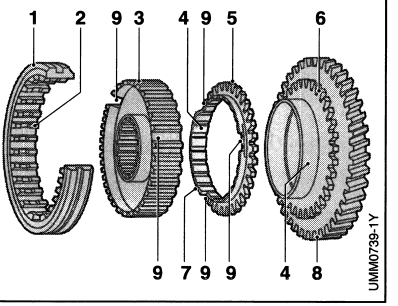
Gear-shift elements

Synchromeshes

During each gear shift or range change in a transmission rotating elements of the target gear are connected to each other. In manually shifted transmissions, automated manual transmissions and dual-clutch transmissions the gear wheel (movable gear) is accelerated or decelerated to the shaft speed. When there is synchronization a frictional connection is established by means of dog toothing. The device that brings about the equal rotational speeds is usually called a synchromesh.

During the synchronizing process a synchronizer ring is pushed via the gearshift sleeve towards a conical friction surface on the gear wheel to be shifted (Figure 5). The process of pressing the synchronizer ring with force F_S onto the conical friction surface of the gear wheel generates by way of sliding friction a torque M_S which would like to match the gear wheel to the speed of the shaft. Lockout toothing on the synchronizer ring prevents (or at least makes it difficult) the dog toothing of the gearshift sleeve from meshing with the shift toothing before the gear wheel and shaft are synchronized (Figure 6). In this process the dog toothing of the gearshift sleeve contacts the tooth angles of the

Figure 5: Synchromesh
1 Gearshift sleeve, 2 Dog toothing,
3 Synchronizer body, 4 Friction surface (conical),
5 Lockout toothing, 6 Shift toothing,
7 Synchronizer ring, 8 Gear wheel,
9 Recesses for sliding blocks (sliding blocks not pictured).



lockout toothing on the synchronizer ring and the friction torque of the synchronizer ring counteracts any evasion of the lockout toothing. Further rotation of the synchronizer ring relative to the gearshift sleeve is prevented by axially movable sliding blocks (not shown in Figure 5).

The recesses for the sliding blocks in the synchronizer body are exactly the same width as the sliding blocks, while the recesses in the synchronizer ring are slightly wider. In this way, the synchronizer ring can, as pictured in Figure 6, move relatively through a small angle. The sliding blocks transmit the torque from the synchronizer body to the synchronizer ring. The external teeth on the synchronizer body, synchronizer ring and shift toothing are dimensioned in such a way that they can establish positive locking with the dog toothing of the gearshift sleeve.

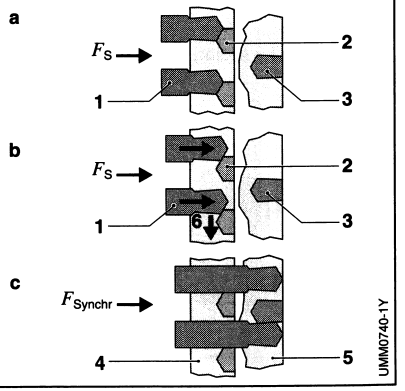
Once the synchronization of the movable gear and the shaft is achieved, the synchronizing torque and consequently also the torque acting on the lockout toothing are zero. The applied shift force F_S now induces, via the tooth angles of the synchronizer-ring lockout toothing, the synchronizer ring to rotate relative to the gearshift sleeve so that the shift-sleeve dog toothing can be slid over the synchronizer-ring lockout toothing into the gear-wheel shift toothing. This establishes positive locking between the shaft and the gear wheel.

Increased shift force can reduce the synchronizing time, but in the long term this will cause increased wear of the friction lining on the synchronizer ring. Also, the lockout toothing can be turned back by increased shift force already before synchronization and the shift-sleeve dog toothing advanced up to the gear-wheel shift toothing. When the shift-sleeve dog teeth and the shift toothing on the gear wheel engage at different speeds, this results in "gear-shift rasping".

Synchromeshes with two to three conical friction surfaces arranged on top of each other are used to synchronize components with higher mass inertia. These are referred to as double-cone or triple-cone synchromesh.

Often two and sometimes even three characteristic forms of synchromesh are fitted in transmissions. For the most part, the gear wheels of first gear and reverse gear are equipped with triple-cone synchromesh, the gear wheels of second and third gears with double-cone synchromesh, and the other gears with single-cone synchromesh. The advantage of this arrangement in manually shifted transmissions is that the shift forces are similarly high for all the gears.

Figure 6: Synchronizing process
Force F_S on the synchronizer ring.
a) Synchronization and lockout,
b) Rotate synchronizer ring back,
c) Establish positive locking.
1 Dog toothing on gearshift sleeve,
2 Lockout toothing on synchronizer ring,
3 Shift toothing on gear wheel,
4 Synchronizer ring, 5 Gear wheel,
6 Relative motion of synchronizer ring in relation to gearshift sleeve.



Multi-disk clutch

A multi-disk clutch usually consists of a number of inner and outer disks. The inner disks have on their inside diameter synchronization toothing for joining the clutch to components in the transmission, while the outer disks have this toothing on their outside diameter (Figure 7). A disk is glued in each friction contact with a paper-like lining (in Figure 7 the disk with internal teeth), the other disk has a smooth steel surface. A higher torque can be transmitted for the same operating force by the use of a number of alternately stacked inner and

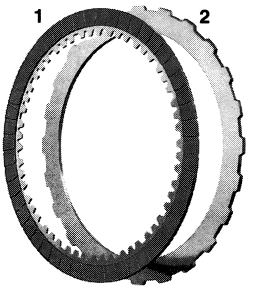
outer disks and the greater lining surface area thereby obtained. To ensure that the multi-disk clutch is efficiently cooled, the paper lining has grooves for a constant throughflow of cooling oil.

A multi-disk brake is a special type of multi-disk clutch. The multi-disk brake differs in that for example the outer disks are connected to a non-rotating component such as the transmission housing and the clutch disks are stationary in the closed state while the elements of a multi-disk clutch rotate in the closed state. The multi-disk brake can be used for example to support a planetary-gear set element against the transmission housing in order to effect a transmission ratio.

The oil-cooled starting clutch and the dual clutch are distinct types of the multi-disk clutch for these load cases.

Multi-disk clutches are of fundamental importance to the function of power-shift transmissions. The gear to be shifted is introduced into the power flow frictionally and without interruption of tractive force [1].

Figure 7: Multi-disk clutch
1 Inner disk,
2 Outer disk.



Manually shifted transmissions

Manually shifted transmissions (MTs) are usually designed as spur-gear drives in which the gear shift and the drive-off process are performed manually by the driver. Normally a dry single clutch is used as the power take-up element. In manually shifted transmissions each gear is realized by a spur-gear pair. The different ratios of the numbers of teeth per spur-gear pair produce the different gear ratios. All the gears are permanently mounted in the transmission. Each gear pair of a gear has a fixed gear which is permanently connected to a shaft and a movable gear which is pivoted on another shaft. For each gear the associated movable gear is non-rotationally connected to the shaft by gear-shift elements (see Gear-shift elements). The individual movable gears are coupled via synchromeshes (see Figure 5).

Manually shifted transmissions for the most part have five to six forward gears and one reverse gear. The spacing (ratio of the gear ratios from the lowest to the highest gear) of the transmissions is roughly four to six.

Manually shifted transmissions belong to the category of tractive-force-interrupted transmissions. To change a transmission step the power flow must be disconnected by the clutch.

The designs for a standard drive (Figure 8 and Figure 9) have a coaxial input and output and a countershaft. A constant transmission ratio transmits the engine power to the countershaft. The movable gears of the different transmission ratios are coupled by synchromeshes to the transmission output.

Types of construction with one (see Figure 10) or two countershafts parallel to the drive have established themselves in front-wheel drives. Transmissions with two shafts parallel to the drive are called triple-shaft transmissions. Splitting the transmission

steps between two parallel shafts enables the transmission to be short in its construction. The overall length can be further optimized for a triple-shaft transmission by the double use of fixed gears on the drive shaft. Here one fixed gear meshes with two movable gears (e.g. for the third and fifth gears) on the different countershafts. In a front-wheel drive the final-drive unit with differential is integrated in the transmission (see Figure 10). The final-drive unit is driven by one (in triple-shaft transmissions by two) constant transmission ratio of the countershafts.

Figure 8: Schematic diagram of a 6-speed manually shifted transmission for a standard drive

1 From driving engine, 2 Dual-mass flywheel, 3 Clutch, 4 Constant transmission ratio (5th gear), 5 Synchromeshes, 6 Movable gear, 7 Countershaft, 8 Fixed gear, 9 To final-drive unit, 10 Transmission input shaft.
I...VI First through sixth gear, R Reverse gear.

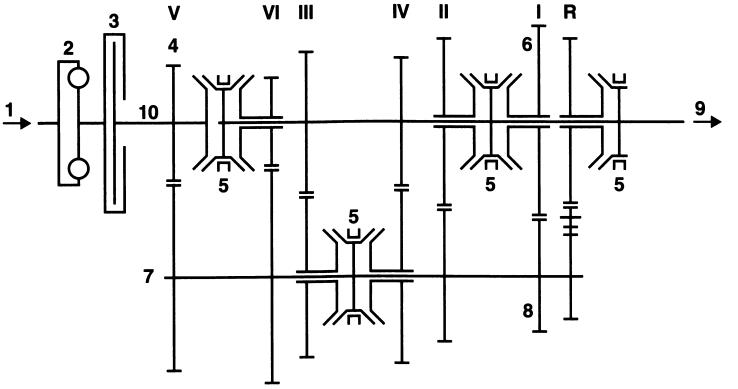
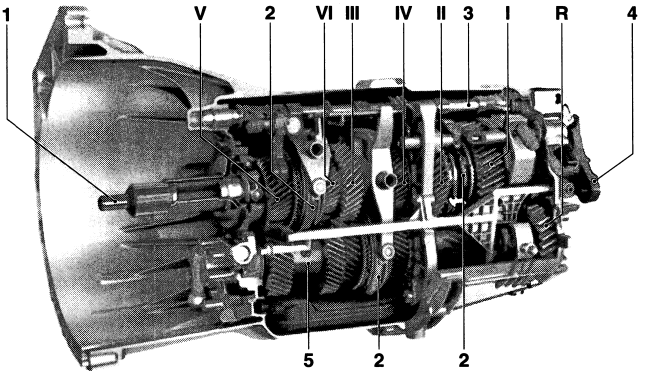


Figure 9: Structural design of a 6-speed manually shifted transmission for a standard drive

1 Transmission input shaft, 2 Synchromeshes, 3 Gearshift linkage, 4 To final-drive unit, 5 Countershaft.
I...VI First through sixth gear,
R Reverse gear.



Automated manual transmissions

Automated manual transmissions (AMTs) have a mechanical design comparable to a manually shifted transmission (Figure 10). The clutch and the synchromeshes are operated by actuators (Figure 11) which are actuated by a transmission control unit. This helps to provide the operating comfort of a fully automatic transmission.

Like the manually shifted transmission, the automated manual transmission is a tractive-force-interrupted transmission, which gives rise to disadvantages in shifting comfort compared with other automatic-transmission construction types.

Figure 11: Structural design of an actuator for an automated manual transmission for front-transverse drive

1 Clutch,
2 Clutch actuator.

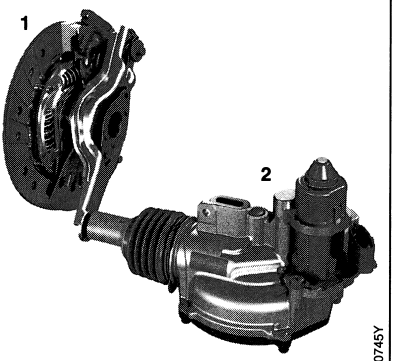
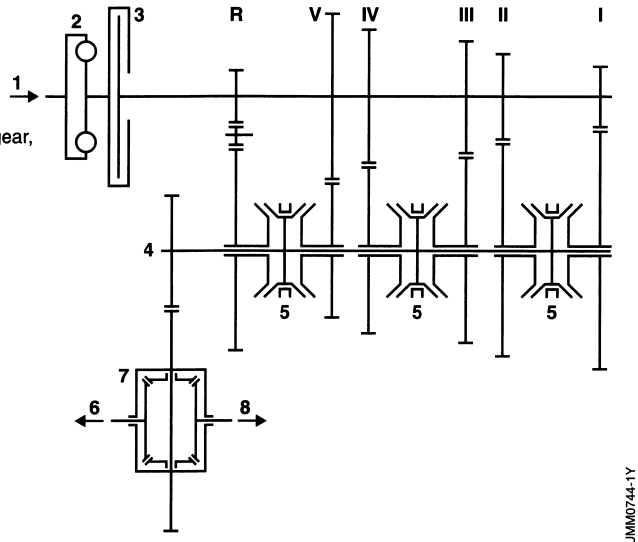


Figure 10: Schematic diagram of an automated manual transmission for front-transverse drive

1 From driving engine,
2 Torsion damper,
3 Clutch,
4 Output constant,
5 Synchromeshes,
6 To wheel,
7 Differential,
8 To wheel.
I...V First through fifth gear,
R Reverse gear.



Dual-clutch transmissions

Dual-clutch transmissions (DCTs) consist of two separate spur-gear sub-transmissions which can be connected in each case via a clutch to the internal-combustion engine [1]. By splitting the even gears (two, four, six) and odd gears (one, three, five, seven) between the two sub-transmissions (Figure 12) it is possible to change the gear ratios by means of power-shifting. With its tractive-force-interruption-free gear shift, a

Figure 12: Schematic diagram of a 7-speed dual-clutch transmission for a standard drive

1 From driving engine, 2 Dual-mass flywheel, 3 Dual clutch, 4 Countershaft,
5 Synchromeshes, 6 To final-drive unit.
I...VI First through seventh gear, R Reverse gear.

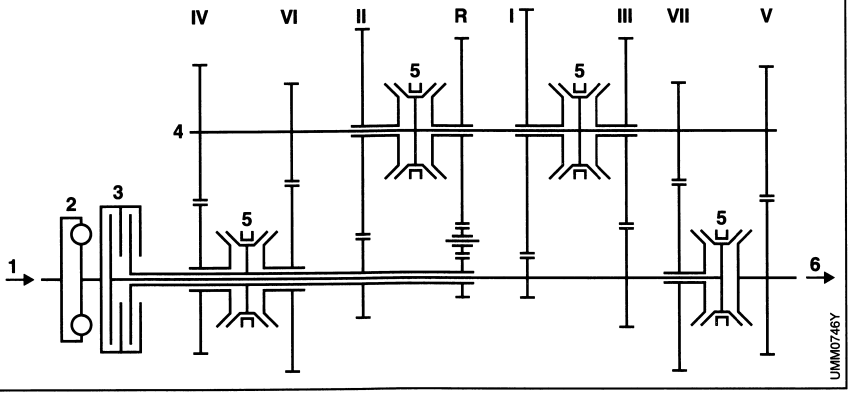
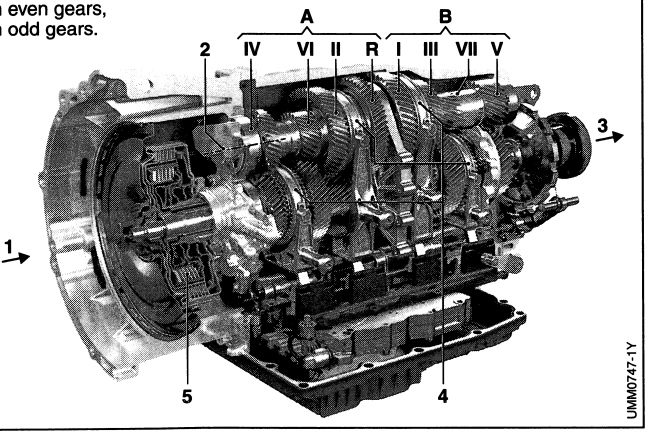


Figure 13: Structural design of a 7-speed dual-clutch transmission for a standard drive

1 From driving engine, 2 Countershaft, 3 To final-drive unit, 4 Synchromeshes,
5 Dual clutch.
I...VII First through seventh gear
(fifth gear as direct ratio),
R Reverse gear,
A Sub-transmission with even gears,
B Sub-transmission with odd gears.



dual-clutch transmission offers the comfort of an automatic transmission. The clutches are used both for load sharing during shifting and for the drive-off process.

The reverse gear can be assigned to either the even or the odd sub-transmission. In Figure 12 the reverse gear is assigned to the even sub-transmission.

The advantage of dual-clutch transmissions is that, unlike automatic planetary transmissions, they are suitable for engine speeds in excess of 7000 rpm. Dual-clutch transmissions (Figure 13) are therefore used among others in sports-car applications.

The spur-gear construction in dual-clutch transmissions facilitates advantageous gear-set versions especially in front-transverse transmissions with their axial offset between transmission input and transmission output shafts. In these gear sets the differential, unlike an automatic planetary transmission, is driven directly by the countershafts, providing for less meshing and consequently greater transmission efficiency in front-transverse transmissions.

Automatic planetary transmissions

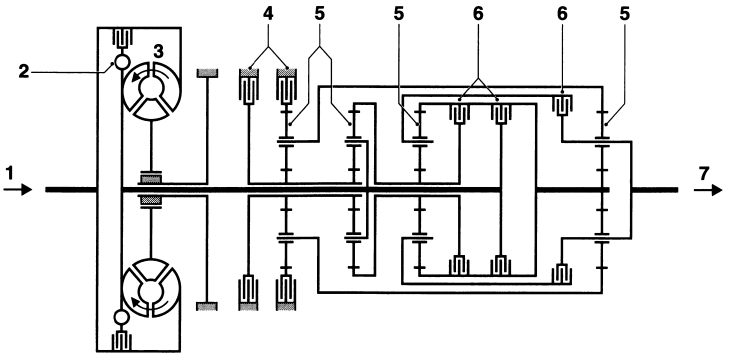
Automatic planetary transmissions (ATs) realize their transmission steps through a series of planetary-gear sets arranged in succession which are connected to each other in different ways by multi-disk clutches (Figure 14). The number and arrangement of planetary-gear sets and clutches are dependent on the number of gears implemented in the transmission.

An 8-speed transmission can for example be achieved with four planetary-gear sets and five multi-disk clutches (Figure 14). By means of different combinations of several simultaneously closed and opened shift elements the transmission schematic shown can realize eight forward gears and one reverse gear. A hydrodynamic torque converter usually serves as the power take-up element.

In automatic planetary transmissions eight to nine gears have proven to be the optimum with regard to reduced fuel consumption and good driveability. For a smaller number of gears the combustion-engine program map is not yet opti-

Figure 14: Schematic diagram of an 8-speed automatic planetary transmission for a standard drive

1 From driving engine, 2 Torsion damper, 3 Hydrodynamic converter, 4 Multi-disk brakes, 5 Planetary-gear sets, 6 Multi-disk clutches, 7 To final-drive unit.



mally developed to drive constantly in the internal-combustion engine's optimum consumption and performance range. For a larger number of gears additional shift elements or planetary-gear sets are required which diminish the efficiency of the transmission and hence the drivetrain in most driving situations. The energy that is required for the greater number of gear shifts can likewise no longer be compensated by better utilization of the combustion-engine program map.

The planetary-gear sets have coaxial input and output. This results in advantageous transmission construction types particularly for the standard drive, where the engine crankshaft and the axle-differential drive are on the same axis (see Figure 15). Since the transmitted power is shared in the planetary-gear sets between a number of planetary gears spread over the circumference, automatic planetary transmissions are characterized by a high power density. The comfort of a power-shift transmission is achieved by targeted overlapping shifts of the clutches.

Continuously variable transmissions

In continuously variable transmissions (CVTs) the transmission-ratio range is achieved by means of a pushbelt or a plate link chain between a first and a second V-pulley set (Figure 16). The pulleys of each V-pulley set are hydraulically pressed on their rotational axis axially towards each other. Axial displacement of the V-pulleys changes the axial clearance between the pulleys and hence the running radii to which the belt element can attach over its unchanging width. The combination of different, infinitely variable running radii on the two V-pulley sets (input and output) produces the gear ratio of the belt variator. Since no change between different gears has to take place, the continuously variable transmission offers high driving comfort without interruptions of tractive force.

Figure 15: Structural design of an 8-speed automatic planetary transmission for a standard drive

1 From driving engine, 2 Torsion damper, 3 Hydrodynamic converter,
4 Multi-disk brakes, 5 Planetary-gear sets, 6 Multi-disk clutches,
7 To final-drive unit.

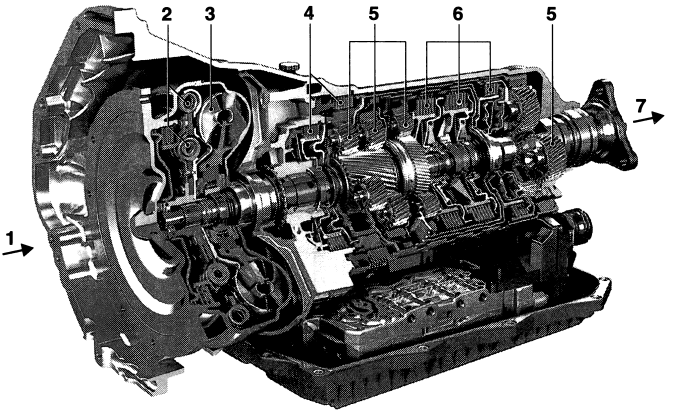
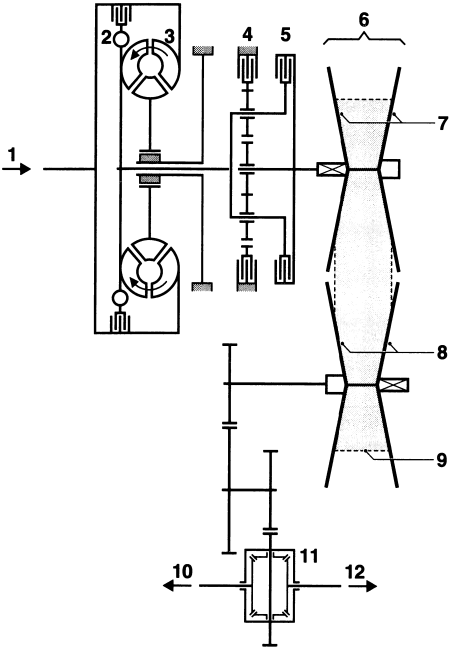


Figure 16: Schematic diagram of a continuously variable transmission

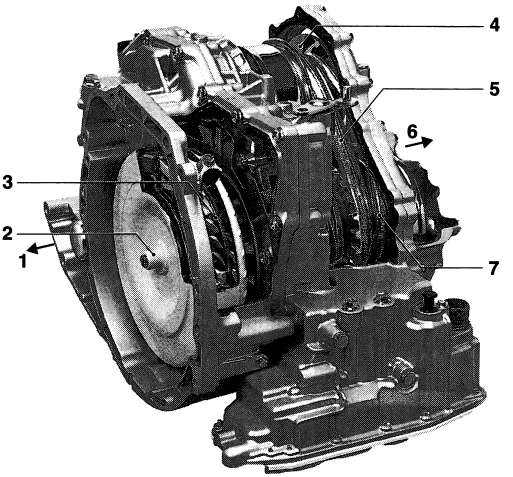
1 From driving engine,
2 Torsion damper,
3 Hydrodynamic converter,
4 Reverse-gear brake,
5 Forward-gear clutch,
6 Belt variator,
7 First V-pulley set,
8 Second V-pulley set,
9 Belt or chain,
10 To wheel,
11 Differential,
12 To wheel.



UMM0750Y

Figure 17: Structural design of a continuously variable transmission (source: Jatco)

1 To wheel,
2 From driving engine,
3 Hydrodynamic converter,
4 Second V-pulley set,
5 Belt,
6 To wheel,
7 First V-pulley set.



UMM0751-1Y

Thanks to the given axial offset between input and output of the belt variator, continuously variable transmissions are ideal for a front-transverse transmission layout (Figure 17). An oil-cooled clutch or a hydrodynamic converter is used as the power take-up element. A planetary-gear set with two clutches is installed to switch between forward and reverse driving (Figure 16).

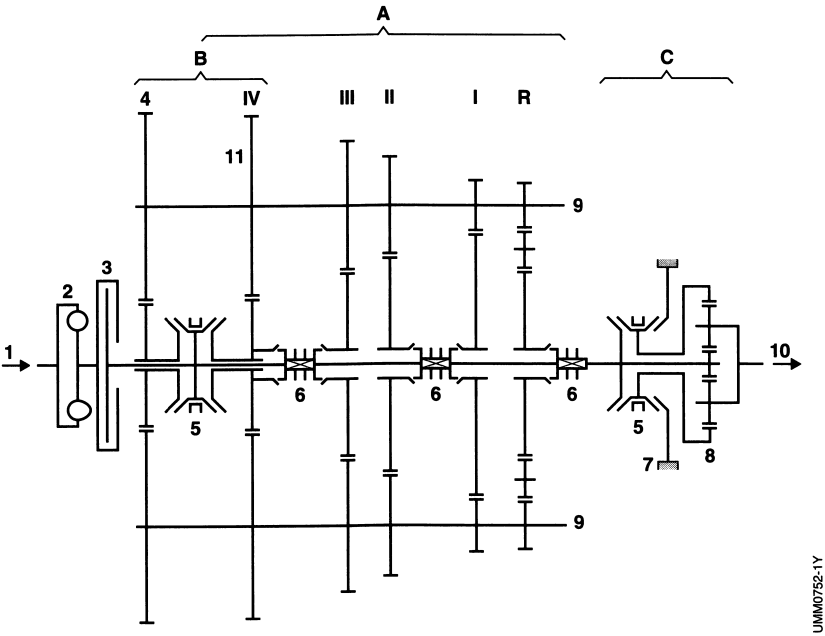
Commercial-vehicle transmissions

The standard construction has gained acceptance as the drivetrain concept in trucks with a gross weight rating of over 3.5 tons. The high loads on the rear powered axle(s) and the associated potential high drive force of the drive wheels are the reason for this drivetrain concept.

The longitudinally or transversally mounted rear engine has gained acceptance in buses in order to guarantee maximum transport space for passengers and luggage.

Figure 18: Schematic diagram of a 16-speed truck range-change transmission with splitter box and range-change box

1 From driving engine, 2 Torsion damper, 3 Clutch, 4 Constant transmission ratio 1,
5 Synchromesh, 6 Shift dogs, 7 Housing connection, 8 Planetary-gear set,
9 Countershafts, 10 To final-drive unit, 11 Constant transmission ratio 2 (fourth gear).
I...IV First through fourth gear in main box.
R Reverse gear.
A Main box, B Splitter box, C Range-change box.



Commercial-vehicle transmissions differ significantly in their dimensions from passenger-car transmissions. The reasons for this are not only higher engine torques and engine power output, but also above all the much higher mileage demands of sometimes over one million kilometers.

The commercial-vehicle sector features predominantly manually shifted transmissions, automated manual transmissions, and planetary converter transmissions. Primarily manually shifted transmissions and automated manual transmissions have become prevalent in trucks. Predominantly planetary converter transmissions and automated manual transmissions are used in buses.

Up to 16 gears are obtained in manually shifted transmissions and automated manual transmissions. To minimize construction expense, two to three sub-transmissions (splitter box, main box, range-change box) are arranged in succession (Figure 18). In a truck transmission with 16 gears the main box has for example four forward gears and one reverse gear. Mounted upstream of this main box is a splitter box with two gears. The splitter box allows intermediate ratios of the four gears of the main box. Often the splitter box is used with a full payload since only small gear steps are possible for the very high vehicle weights in relation to the drive power. When the truck is unladen, the ratio of engine power to vehicle weight is much more favorable such that the greater gear steps of the main box can be shifted without using the splitter box.

Splitter and main boxes transmit the torque in the transmission variant shown to two countershafts. This enables narrower gear widths for the very high torques in the truck and a smaller center distance between the shafts to be chosen.

Mounted downstream of the main box is a range-change box. The transmission step between the two gears of the range-change box corresponds to the ratio spacing of the main box from first to fourth gear. This means that when shifting the range-change box gear from slow to fast gear at the same time the main box is shifted from fourth to first gear and vice versa. In this way, the spacing of the main box is utilized twice. In this way, multiplying two splitter gears by four gears of the main box and two gears of the range-change box produces the maximum number of gears for these transmissions of 16 possible forward gears.

The planetary converter transmissions used predominantly in buses as a rule have six gears. These transmissions have essentially the same as those in a passenger car, but their dimensions are adapted to the higher torque and mileage demands.

Electronic transmission control

Function and requirements

Today, primarily electronically actuated hydraulic systems are used to control automated transmissions. Hydraulic actuation is retained for the clutches, while the electronics assume responsibility for gear selection, for gearshifting, and for adapting the pressures in accordance with the torque flow. The advantages are:

- A number of gearshift programs including adaptive drive functions,
- Improved, constant gearshift comfort over the full service life,
- Adaptability to different types of vehicle,
- Simplified hydraulic control.

Alternatively, there are electromechanically actuated transmissions, where electric motors assume the role of the hydraulics (actuating clutch, engaging and selecting gear).

Sensors detect the transmission output- and input-shaft speeds, the pressure at the clutches, the selector-lever position, and the positions of the program selector and the kick-down switches. The information about the engine (speed, driver command, and engine torque) is transmitted by the engine ECU via the CAN bus to the transmission control. The ECU processes this information according to a predefined program and uses the results in determining the control variables to be transmitted to the gearbox. Electrohydraulic converter elements form the link between the electronic and hydraulic circuits. Pressure regulators ensure precise control of pressure levels at the clutches.

Table 1: Transmission control for the different transmission types

	AT	CVT	DCT
No. of actuators	<10	<6	<12
Pressure range [bar]	<20	<70	< 30
No. of sensors	4...6	4...6	8...12
Computing power	High	Medium	Very high

AT Automatic Transmission
CVT Continuously Variable Transmission
DCT Dual-Clutch Transmission

The transmission control must satisfy different requirements with regard to the transmission types and thus also varies greatly in its circuit complexity. This ranges from the simple remote-mounted ECU (where the ECU is separated from the transmission) for a continuously variable transmission (CVT) to the complex electronic module for a dual-clutch transmission (Table 1).

Transmission-control components

Selector lever

The classic selector lever in the center console with the positions P, R, N and D and a cable to the transmission is replaced in today's shift-by-wire system by simple rotary or rocker switches which permit more freedom in the vehicle's passenger-compartment design. However, the main functions have been retained and are the same for all automated transmissions:

- P Park; the parking lock is activated in the transmission, preventing the vehicle from rolling.
- R Reverse engages the reverse gear.
- N Neutral or idle; i.e. there is no power transmission between the engine and transmission.
- D Drive; forward driving is activated.

Many systems still have a so-called M gutter which allows the gear to be manually selected.

Program selector

The driver can select the following programs by pushbutton or switch:

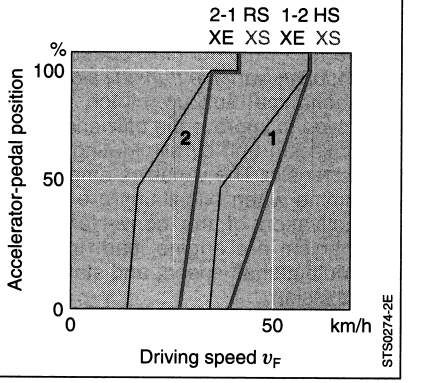
- E Economy, economical driving in as high a gear as possible (Economy mode).
- S Sport program, performance-oriented gear selection (Sport mode).
- M Manual gear selection.

Selection of the Economy mode or the Sport mode is very often performed by intelligent shift-point control in the software.

Shift-point control
In selecting the gear to be engaged, the system refers to the rotating speed of the transmission output shaft and the driver command before triggering the appropriate solenoid valves or pressure regulators. Either the driver chooses between different driving programs (e.g. for optimum consumption- or performance-based driving) or an adaptive driving program is implemented which selects the optimum gear from the driving situation and driver type. In addition, the selector lever allows manual input from the driver. Figure 1 shows the upshift curve for shifting from first to second gear in Economy and in Sport modes, i.e. the accelerator-pedal position and speed at which the gear shift takes place. The same applies to downshifting. There must be a gap between the upshift and downshift curves to avoid gear hunting.

Driving-situation detection includes variables such as linear and lateral acceleration, and results in the detection of hills or corners. For this purpose progressive systems additionally incorporate the information from the navigation system in order already to prepare future driving situations. The speed with which the accelerator pedal is pressed and the number of kick-down operations in a defined period of time or the frequency of brake-pedal operation serve to identify the type of driver (from very economy- to very performance-minded). A complex control program selects the appropriate gear for the current operating conditions and driving style. In other words, it suppresses trailing-throttle upshifts on the approach to and during corners, or automatically responds to slow throttle openings by activating a drive program for low engine-speed upshifts.

Figure 1: Upshift and downshift curves in Economy and Sport modes
1 Upshift curves,
2 Downshift curves.
XE Economy mode, XS Sport mode,
DS Downshift, US Upshift.
Range above 100 % is triggered by kick-down switch.



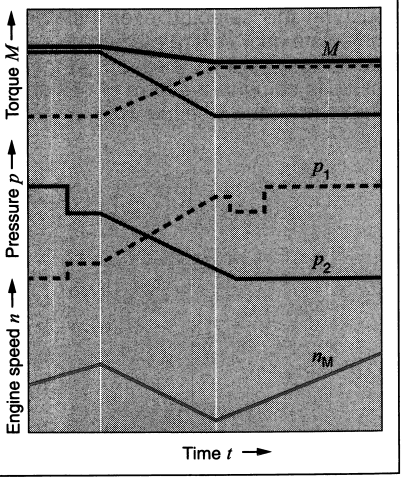
Concepts which combine the high level of convenience of such "intelligent" gear-shift programs with facilities for active adaptation to individual driver preferences have become very widespread. The selector levers of these systems have a second parallel gutter (M gutter), in which simply touching the selector lever initiates an immediate gear shift (provided no engine-speed limits are exceeded). Alternative solutions come in the form of paddles or ± buttons on the steering wheel.

Gearshift sequence from the example of an upshift in a multi-speed transmission
Power upshifts take place in an automatic transmission, unlike a manual transmission, without interruption of tractive force as overlap control. While clutch 1 opens for gear x, clutch 2 must close for gear y (Figure 2).

Converter lockup (for multi-speed and continuously variable transmissions)
A mechanical lockup clutch can be employed to improve the efficiency of the transmission unit by eliminating torque-converter slip. The variables employed to determine when conditions are suitable for activation of the converter lockup mechanism are engine load, transmission output-shaft speed, and state of the transmission.

Control of shift quality
The accuracy with which the pressure at the friction elements of the clutches is adjusted to the level of torque being transmitted (determined with reference to engine load and speed) has a decisive influence on shift quality; this pressure is regulated by a pressure regulator. Shifting comfort can be further enhanced by briefly reducing engine torque (via an engine intervention) for the duration of the shift (e.g. in a gasoline engine by retarding the ignition timing, reducing the injected fuel quantity, or cylinder cutout); this also reduces friction loss at the clutches and extends component service life.

Figure 2: Overlap control
p_1 Pressure, clutch to be cut in,
p_2 Pressure, clutch to be cut out,
n_M Engine speed,
M Torque.



Adaptive pressure control
Crucial parameters such as clutch fill time and slip time are permanently monitored to ensure a constant gearshift quality throughout the transmission's service life. In the event of deviations from specified limits, pressure corrections are made in the system in order to maintain an optimum time characteristic during the next gearshift and thus optimum shift quality.

Safety strategy
Special monitoring circuits and functions prevent transmission damage stemming from operator error, while the system responds to malfunctions in the electrical system by reverting to a backup mode. It is important to ensure at all times that vehicle instability does not arise.

Software-development requirements
Software development for automated transmissions must satisfy the following requirements:
- Compliance with functional safety (ASIL D, Automotive Safety Integrity Level, [2]),
- ASPICE (Automotive Software Process Improvement and Capability Determination) as the development process in the international development network [3],
- AUTOSAR as the interface standard (Automotive Open Systems Architecture, see AUTOSAR), and
- software sharing between ECU manufacturers, transmission manufacturers, and vehicle manufacturers.

Hydraulic control

Functions

The main purpose of hydraulic control is to regulate, boost and distribute hydraulic pressures and volumetric flows. This includes generating the clutch pressures, supplying the converter, and providing the lubricating pressure. The hydraulic-control housings are made of diecast aluminum and contain a number of finish-machined slide valves and electrohydraulic actuators. Figure 3 shows the hydraulic module of an electronic transmission control.

Electrohydraulic actuators

Electrohydraulic converter elements such as solenoid valves and pressure regulators form the interface between electronic and hydraulic circuits. A distinction is also made here between pilot-controlled and directly controlled systems. In pilot-control systems (Figure 4) the pressure regulators operate in a low pressure range (< 10 bar) and additional boost slide valves are needed to achieve the clutch adjustment pressure (e.g. approximately 20 to 25 bar in the multi-speed transmission). In the case of direct control (Figure 5) the pressure regulator can control the entire working-pressure range and requires no additional hydraulic boosting.

Electromechanical control

Electric motors are used as actuators in semi-automatic manual transmissions and some dual-clutch transmissions with electromechanical control. These are designed for dynamic and wear reasons as BLDC (Brushless Direct Current) motors.

References

[1] H. Naunheimer; B. Bertsche; G. Lechner: Fahrzeuggetriebe – Grundlagen, Auswahl, Auslegung und Konstruktion. 2nd Ed., Springer-Verlag, 2007.
[2] ISO 26262-9: Road vehicles – Functional safety – Part 9: Automotive Safety Integrity Level (ASIL)-oriented and safety-oriented analyses.
[3] http://vda-qmc.de/software-prozesse/automotive-spice/

Figure 3: Hydraulic module of a transmission control
1 Transmission connector (to wiring harness),
2 Pressure sensors,
3 ECU,
4 Connectors for contacting the hydraulic valves,
5 Sensor domes of position sensors (rear side),
6 Sensor domes of speed sensors (rear side).

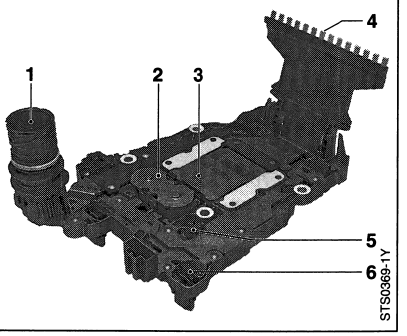


Figure 4: Pilot control (principle)
1 Hydraulic-oil feed to actuator,
2 Oil sump,
3 Feed to slide valve,
4 Actuator,
5 Slide valve in control housing,
6 Clutch.
The arrows indicate the oil flow.

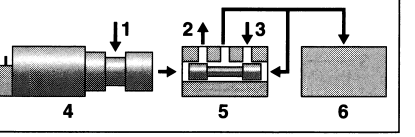
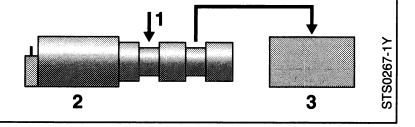


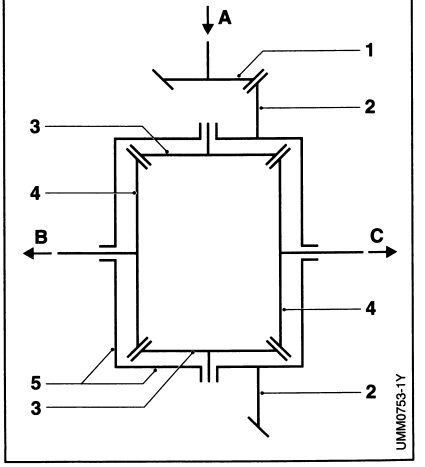
Figure 5: Direct control (principle)
1 Hydraulic-oil feed to actuator,
2 Actuator,
3 Clutch.
The arrow indicates the oil flow.



Differentials

A differential has one input and two outputs. Differentials are most commonly used as final-drive differentials. Their purpose is to distribute uniformly the torque to the two drive wheels when driving straight ahead and when cornering. In all-wheel drive vehicles differentials are used as transfer cases to distribute torque between the vehicle axles.

Figure 1: Schematic diagram of a bevel-gear differential
A Input, propeller shaft,
B Output, drive shaft, left,
C Output, drive shaft, right.
1 Bevel-gear pinion,
2 Bevel crown wheel,
3 Differential bevel gears,
4 Drive-shaft gears,
5 Differential case.



Final-drive differential

The final-drive differential distributes the torque coming from the manual transmission uniformly to the two drive shafts of a vehicle axle. Since when cornering the inner cornering wheel describes a smaller circle and thus turns more slowly than the outer cornering wheel, the final-drive differential must be able to compensate differential speeds between the drive wheels on an axle.

Bevel-gear differential
On a standard drive the input and output of the bevel-gear differential are offset by 90°. The bevel-gear differential drive transmits the torque from the bevel-gear pinion to the bevel crown wheel (Figure 1 and Figure 2). The bevel crown wheel (differential drive gear) transmits the torque via the differential case to the axle shafts of the differential bevel gears. The differential bevel gears distributes the torque to the two drive-shaft gears.

When the vehicle is driving straight ahead the two drive-shaft gears rotate at the same speed. The differential bevel gears circulate at the speed of the bevel crown wheel without rotating about their own axle shaft. They act as drivers and distribute the torque uniformly to the two drive shafts.

When cornering the differential bevel gears compensate the difference in speed of the drive shafts by rolling off on the drive-shaft gears rotating at different speeds. In this process the differential bevel gears rotate on their axle shafts, which are mounted in the differential case.

In most final-drive units the bevel-gear pinion meshes outside the axis center of the bevel crown wheel (Figure 3). This is called a hypoid drive. The advantages of a hypoid drive are:
- Smoother running since a greater number of gears mesh with each other.
- Higher torque transmittability since the bevel-gear pinion has a larger effective diameter due to the axial offset.

The drawback of this arrangement is the increase in sliding as opposed to rolling-off

between the bevel-gear pinion and the bevel crown wheel. However, this can be successfully countered with particularly pressure- and shear-resistant hypoid gear oils.

In front-transverse transmissions the bevel-gear pinion and the bevel crown wheel are replaced by two spur gears acting as pinion and differential drive gear (see Automated manual transmissions). Since the differential gears are designed as bevel gears, this design type is still classed as a bevel-gear differential.

Spur-gear differential

In a spur-gear differential the differential gears are designed as spur gears. There are various design types which utilize the different connection forms of planetary transmissions. In the differential pictured in Figure 4 the torque is routed via the differential drive gear directly into the planetary-gear carrier. The planetary-gear carrier distributes the torque to the planetary gears arranged in pairs on both sides. The planetary gears of a pair mesh with each other in each case with half the face width. The outer halves of the planetary gears mesh with the sun gears. The sun gears direct the torque in each case to the drive shafts [1]. Unlike a standard planetary transmission, there is no ring gear here. The differential case is the planetary-gear carrier for both planetary-gear sets. Sun gear 1 is at the same time drive-shaft gear 1, sun gear 2 is drive-shaft gear 2. Planetary gears 2 in Figure 4 are drawn with dashed lines because they are offset to the rear in the leaf plane and interlocked with planetary gears 1 with half the face width.

When the vehicle is driving straight ahead the two drive shafts rotate at the

Figure 3: Hypoid drive

1 Bevel-gear pinion,
2 Bevel crown wheel,
3 Axial offset.

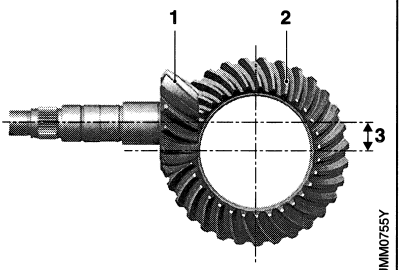
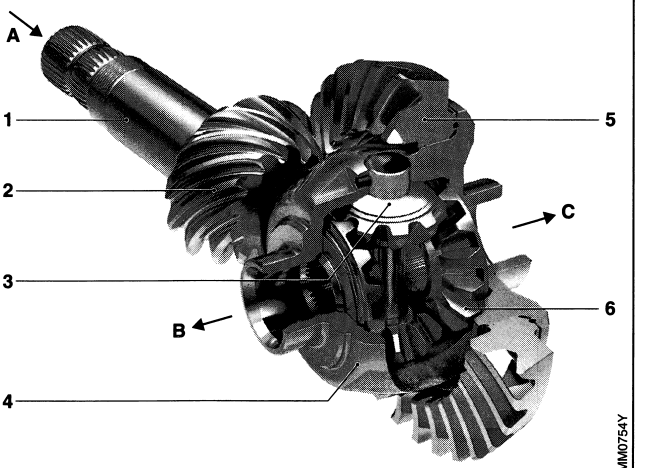


Figure 2: Structural design of a bevel-gear differential

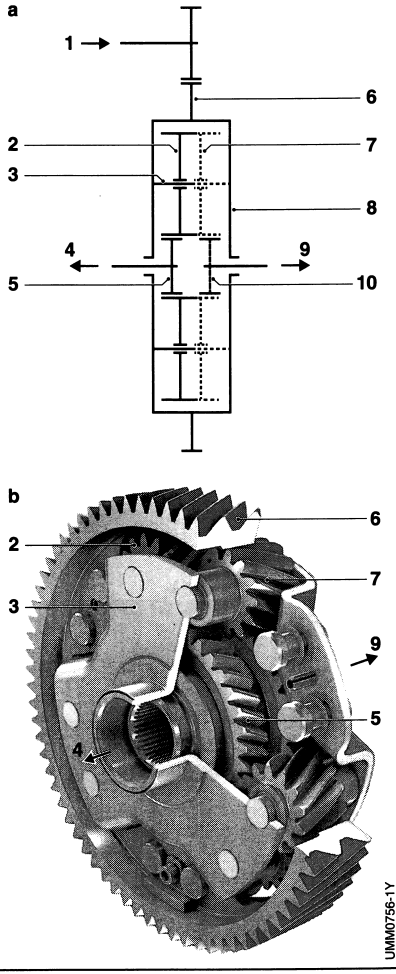
1 Drive shaft, bevel-gear pinion,
2 Bevel-gear pinion,
3 Differential bevel gear,
4 Differential case,
5 Bevel crown wheel,
6 Drive-shaft gear.
A Input, propeller shaft,
B Output,
drive shaft, left,
C Output,
drive shaft, right.



same speed. The planetary gears circulate at the speed of the differential drive gear without rotating about their own axle shaft. They act as drivers and distribute the torque uniformly to the two drive-shaft gears.

When cornering the planetary gears compensate the difference in speed of the drive shafts by rolling off on the sun gears rotating at different speeds. In this process the planetary gears rotate on their axle shafts in the planetary-gear carrier.

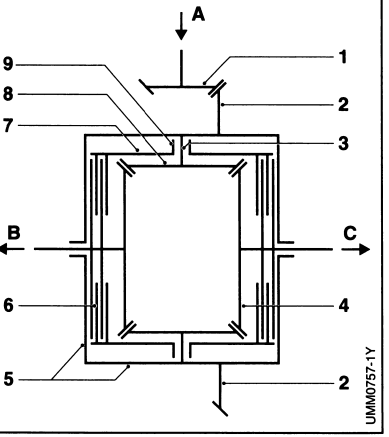
Figure 4: Spur-gear differential
a) Schematic diagram,
b) Structural design.
1 Input, 2 Planetary gear 1,
3 Planetary-gear carrier,
4 Output, drive shaft, left,
5 Sun gear 1, 6 Differential drive gear,
7 Planetary gear 2, 8 Differential case,
9 Output, drive shaft, right, 10 Sun gear 2.



Locking differential

A differential can transmit torque only as long as both vehicle wheels can transmit this torque to the roadway. As soon as the adhesion at a wheel breaks down (e.g. due to black ice), the wheel begins to spin. Since rotational speed but not torque is transmitted on the drive shaft of this wheel, the differential gears cannot be supported on the drive-shaft gear of the spinning wheel and torque is transmitted to the other drive-shaft gear. Consequently, the other wheel stops and transmits no drive torque, although adhesion with the roadway would be possible at this wheel in some cases.

Figure 5: Schematic diagram of a multi-disk self-locking differential
1 Bevel-gear pinion, 2 Bevel crown wheel,
3 Axle shaft, differential bevel gear,
4 Drive-shaft gear, 5 Differential case,
6 Multi-disk clutch,
7 Thrust rings, 8 Differential bevel gear,
9 Edge, thrust ring.
A Input, propeller shaft,
B Output, drive shaft, left,
C Output, drive shaft, right.



The purpose of the locking differential is to transmit drive torque to the two drive shafts, even if a wheel of a driver shaft has reduced adhesion with the roadway.

The locking differential has, in addition to the components described above, multi-disk clutches installed between the differential case and the driver shafts (Figure 5). One thrust ring per multi-disk clutch is installed in the differential case. The thrust rings are non-rotationally connected to the differential case but can move axially to be able to press against the multi-disk clutches. The axle shafts of the differential gears are mounted in recesses between the thrust rings (Figure 6). The axial forces that are produced during power transmission with bevel gears press the thrust rings onto the multi-disk clutches, and this is done more intensely the more torque is transmitted in the bevel-gear drive. The higher the transmitted torque, the higher the locking torque of the multi-disk clutch can be.

With equally good adhesion at both wheels, the majority of the torque is transmitted via the thrust rings to the differential gears, which distribute the torque – since they function as drivers – uniformly to the two drive shafts (Figure 5).

If, due to the loss of adhesion at a wheel, a difference in speed arises between the two drive shafts and thus also between the differential case and the drive shafts, torque is transmitted by the pressed multi-disk clutch to the drive shaft that is rotating at the slower speed.

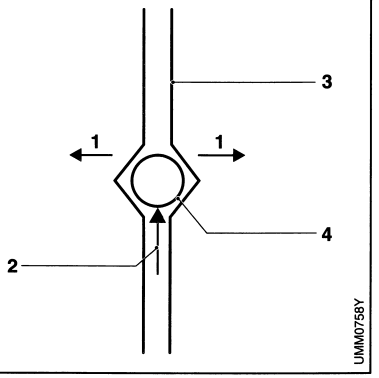
The level of torque and hence the degree of differential locking are dependent on the number of disks and their coefficient of friction, which can be varied by way of different friction linings.

Beyond this design variant there are further solutions approaches for locking differentials, e.g. an actively activatable multi-disk clutch between the drive shafts.

Figure 6: Schematic diagram of thrust rings and axle shafts of differential bevel gears of a multi-disk self-locking differential

The cutting plane is perpendicular to the cutting plane in Figure 5.

1 Spreading force, thrust rings, simultaneously contact pressure, multi-disk clutch,
2 Application of force from differential case,
3 Edge, thrust rings,
4 Axle shaft, differential bevel gear (top view).



Brake intervention with locking-differential effect

Since the introduction of antilock braking systems and driving-dynamics control (Electronic Stability Program) it has been possible to activate brakes of individual wheels. This opportunity is used to brake the faster wheel starting from a specific difference in speed between two drive shafts. Here, torque from the drive shaft rotating at the faster speed in relation to the body is supported via the brake. This enables the differential gears to transmit torque to the drive shaft rotating at the slower speed even without a lock in the differential.

References

[1] B.-R. Höhn; K. Michaelis; M. Heizenröther: Kompaktes Achsgetriebe für Fahrzeuge mit Frontantrieb und quer eingebautem Motor. ATZ – Automobiltechnische Zeitschrift Edition 1/2006.

Drivetrain with hybrid drives

Features

A hybrid vehicle has two or more energy accumulators with their associated energy converters which drive the vehicle jointly or separately (EU Directive 2007/46/EC, [1]). In most hybrid vehicles an electric machine operates as well as the internal-combustion engine. The energy for driving is drawn from a fuel tank and a traction battery. The traction battery is based for the most part on lithium-ion technology with voltage levels of 48 to 800 V. Nickel-metal hydride rechargeable batteries are still used in first-generation hybrid vehicles.

Hybrid vehicles combine the benefits of both types of drive in the drivetrain. The extended ranges of an internal-combustion engine are combined with the high efficiency of an electric machine in urban environments at low driving speeds and with many stops at traffic lights. Essentially three objectives are pursued with the use of hybrid drives:
- reduced fuel consumption,
- reduced pollutant emissions, and
- increased torque and power output to improve operative dynamics.

The internal-combustion engine and the electric machine can appear in the drivetrain in a variety of architectures. Here the vehicle transmission – staged or continuously variable – also plays a role, mostly to provide the internal-combustion engine with variable gear ratios for power conversion.

The electric machine offers high torques at low engine speeds. It thus complements the internal-combustion engine, whose torque only starts to increase at medium engine speeds. The electric machine and the internal-combustion engine contribute to vehicle motion to varying degrees, depending on the drive-power demanded.

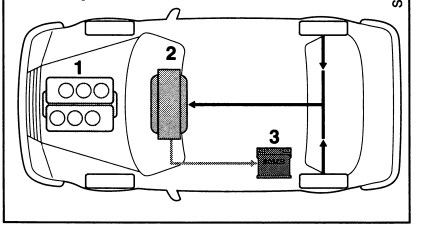
Functionalities of hybrid drives

The way in which internal-combustion engine, electric machine and battery interact determines the possible hybrid functionalities. Hybrid-drive systems can be classified on the basis of the implemented functions and the proportion of electric power that goes towards the entire drive (Table 1).

Start/stop function
The start/stop function switches the internal-combustion engine off temporarily without the driver having to turn off the ignition. The engine is typically switched off when the vehicle is stationary. It is then restarted automatically as soon as the driver is ready to resume driving.

Recuperation
In the case of recuperation (energy recovery) the electric machine is used to convert kinetic energy into stored electrical energy. Braking operations are effected where possible not – or at least not only – by the friction torque of the service brake, but by a generative braking torque of the electric machine. The electrical energy is fed to the battery (Figure 1). The braking process is therefore performed in part in an energy-conserving way.

Figure 1: Recuperation
The arrows indicate the energy flow.
1 IC engine,
2 Electric machine,
3 Battery.



Hybrid driving

Hybrid driving refers to those states in which the internal-combustion engine and the electric machine deliver the desired drive torque together. The assistance provided by the electric machine makes it possible to operate the internal-combustion engine by means of operating-point optimization predominantly in the range of its best efficiency or in operating ranges in which only low pollutant emissions occur. Hybrid driving can be subdivided into generator and motor modes of the electric machine.

Generator mode

In generator mode the internal-combustion engine is operated in such a way that it delivers a higher power output than is needed for the desired propulsion of the vehicle (Figure 2). The excess combustion-engine power is supplied to the electric machine and converted into electrical energy which is used to charge the battery.

Motor mode

In motor mode, also called boosting, the electric machine supports the internal-combustion engine in providing the desired drive power. The battery is discharged in this operating state (Figure 3).

Two different objectives are pursued with boosting: Vehicle performance can be increased compared with a purely engine-powered drivetrain. The second possibility is to maintain the existing overall drive power and use a smaller, more efficient internal-combustion engine. This possibility is called power-neutral downsizing.

Purely electric driving

In the case of purely electric driving the vehicle is driven solely via the electric machine. For this operating mode the internal-combustion engine is disengaged from the vehicle drive and switched off. With the electric drive the vehicle drives quietly, without local emissions and fuel-efficiently on the energy from the battery (Figure 4).

Figure 2: Hybrid driving, generator mode
The arrows indicate the energy flow.
1 IC engine,
2 Electric machine,
3 Battery.

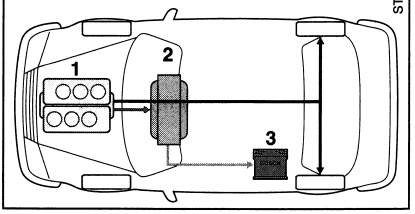


Figure 3: Hybrid driving, motor mode
The arrows indicate the energy flow.
1 IC engine,
2 Electric machine,
3 Battery.

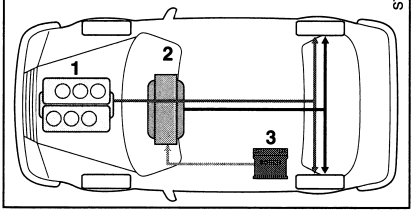


Table 1: Classification of hybrid drives

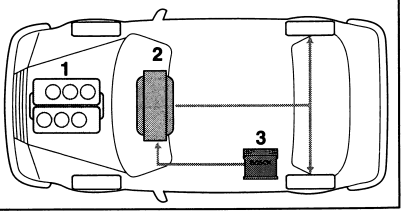
		Hybrid drive			
		Autarkic hybrid			
		Start/stop system	Mild hybrid	Full hybrid	Plug-in hybrid
Function	Start/stop function	●	●	●	●
	Recuperation	●	●	●	●
	Hybrid driving		●	●	●
	Purely electric driving			●	●
	Charging from the mains-supply network				●

Charging from the mains-supply network

Charging the battery from the mains-supply network is performed in the simplest case using a cable connection and a charger which is supplied from a power socket with alternating current or from a three-phase AC connection with energy. The limited power afforded by common household power supplies necessitates charging times of several hours already for ranges below 50 km. Rapid-charging stations operate with direct current and reduce the charging time significantly.

As an alternative to the plug-in connection, the battery can be charged inductively. Electric alternating fields and transformers are used here. With a primary-side field coil on the ground which is supplied from the mains-supply network electrical energy is transmitted to a vehicle-side induction coil and converted by a charger installed in the vehicle into direct current to charge the battery.

Figure 4: Purely electric driving
The arrows indicate the energy flow.
1 IC engine,
2 Electric machine,
3 Battery.



Classification based on proportion of electric power

Hybrid drives can be classified according to the amount of power the electric drive contributes to the overall power of the vehicle and the possible operating states (see Table 1).

Start/stop system

A start/stop system implements the start/stop (see start/stop function) and recuperation functions. During recuperation the alternator increases in this system its power in the vehicle's overrun mode (with zero fuel consumption) to convert as much kinetic energy as possible into starter-battery electrical energy (12 V). During vehicle accelerations and at constant driving speeds the alternator decreases its power depending on the state of battery charge as far as possible to reduce the combustion-engine load caused by the alternator and thereby reduce the fuel consumption.

There is no battery for electric driving.

Mild hybrid

With a mild-hybrid system the hybrid-driving function can be implemented in addition to the functions of a start/stop system. An electric machine assists the internal-combustion engine here.

As well as the 12V battery, the mild hybrid contains a traction battery with a higher voltage level (e.g. 48 V) from which the electric machine draws its energy.

Full hybrid

A full-hybrid system can, in addition to the functions of a mild-hybrid system, drive over shorter distances with the electric machine alone. The internal-combustion engine is disengaged during purely electric driving from the drivetrain by the power take-up element or an interrupting clutch and switched off (see Figure 5c). The traction battery serves first and foremost to absorb the energy during recuperation or with excess combustion-engine power in the case of operating-point shifts. The voltage level of the traction batteries in a full hybrid can be up to 800 V. The traction batteries used are not sufficient to facilitate purely electric driving over longer distances.

Because the battery is not rechargeable from an external source, these vehicles are from the electric-storage perspective "autarkic". The full hybrid is therefore assigned to the group of autarkic hybrid systems (see Table 1).

Plug-in hybrid
An extension of the full-hybrid system is the plug-in hybrid, whose traction battery is not charged exclusively by the internal-combustion engine or by recuperation, but can also be charged from the mains-supply network. This technical aspect is promoted by legislation, e.g. in Europe, the USA, but also in China, since here the purely electric driving function is taken into account based on range in the CO_2 certification. Such drives with larger traction batteries that can be recharged from an external source are, to reflect the charging plug used, called plug-in hybrid drives. To better utilize the available electrical energy, they are as a rule equipped with larger electric motors than autarkic hybrids.

In principle plug-in hybrids constitute a bridge technology between autarkic hybrid vehicles and purely battery-operated electric vehicles. The benefits of both systems can be combined here. A plug-in hybrid vehicle is able, depending on the configuration, to cover small to medium distances purely electrically and thus without local emissions (e.g. typical commuter distances) and by using the internal-combustion engine attain its maximum system performance and cover extended ranges [2].

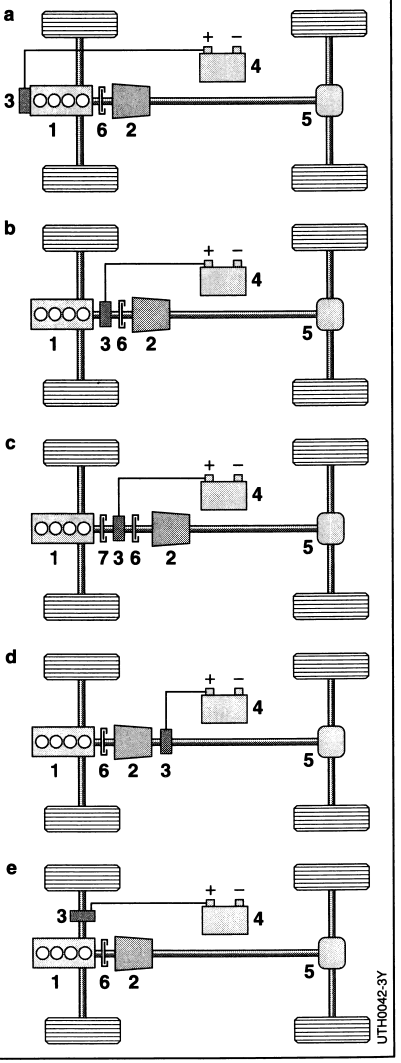
Classification based on drive structure

The different hybrid drives can be classified on the basis of how the electric machine is integrated in the drivetrain.

Parallel hybrid
In a parallel hybrid drive both the internal-combustion engine and the electric machine influence the vehicle drive simultaneously (parallel). The electric machine of the parallel hybrid drive can basically be positioned at different points in the drivetrain, which brings with it specific advantages and disadvantages. A nomenclature originally defined by Daimler AG has gained acceptance for designation purposes; this system designates the parallel-hybrid drivetrain based on the position of the electric machine in the drivetrain as P0 through P4. The P here stands for the parallel architecture, while the number denotes the installation location of the electric machine in the drivetrain. Figure 5 shows this arrangement of the electric machine in the example of a standard drive. This also applies to the front-transverse drive.

The advantage of the parallel hybrid is that for the hybrid versions of the respective vehicle no new transmission has to be developed, but instead a modular construction kit can be used to realize both conventional transmissions (dual-clutch or automatic planetary transmissions as in Figure 8, but also mechanical CVTs) only in conjunction with internal-combustion engine and also hybrid transmissions with different power stages of the electric machine and different traction-battery sizes.

Figure 5: Schematic diagram of a parallel hybrid architecture with various arrangements of the electric machine in a standard drive.
a) P0, b) P1,
c) P2, d) P3,
e) P4 (final-drive unit).
1 IC engine, 2 Manual transmission,
3 Electric machine, 4 Traction battery,
5 Final-drive differential, 6 Power take-up element, 7 Interrupting clutch.



Axle hybrid
The axle hybrid (P4) can be considered a special form of parallel hybrid. In this case, one of the two vehicle axles is driven by an internal-combustion engine with a conventional transmission connected downstream. The second vehicle axle is driven by an electric machine (Figure 6). This concept delivers, aside from the above-mentioned hybrid functions, a four-wheel-drive functionality limited by the power of the electric machine and battery charge.

The transmission for the internal-combustion engine is the same as that of a vehicle without axle hybrid. Particularly for front-transverse drives with very limited installation space it can be advantageous not to have to accommodate the electric machine between the engine and the transmission. The downside of this concept is that the electric machine may not be able to utilize the different gears of the combustion-engine transmission and is subject to losses at high driving speeds as long as it cannot be disengaged from the final-drive differential and the wheels.

Series hybrid

In a series-hybrid drive there is no mechanical connection of the internal-combustion engine to the vehicle drive. The internal-combustion engine generates by means of a generator electrical energy, which is converted in a second electric machine working as the drive back into mechanical energy (Fig-

Figure 6: Schematic diagram of an axle hybrid (P4) in a front-transverse drive
1 IC engine, 2 Manual transmission,
3 Final-drive differential, 4 Drive shaft,
5 Traction battery, 6 Electric machine,
7 Step-up gear unit (fixed ratio).

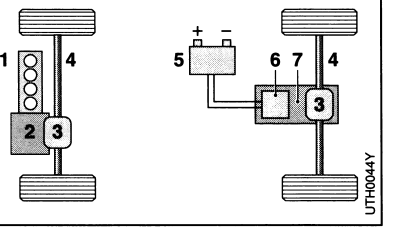
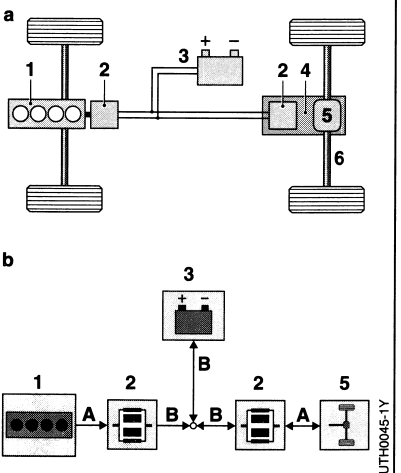


Figure 7: Series-hybrid architecture
a) Schematic diagram of drivetrain,
b) Energy flow.
The arrows indicate the direction of the energy flow.
1 IC engine,
2 Electric machine, 3 Traction battery,
4 Step-up gear unit (fixed ratio),
5 Final-drive differential, 6 Drive shaft.
A Transmission of kinetic energy,
B Transmission of electrical energy.



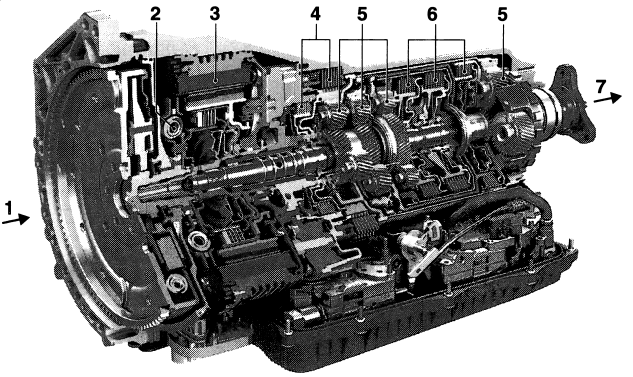
ure 7). Standing against the high driving comfort provided by the electric machine without ratio change and the possibility of operating the internal-combustion engine at its optimum consumption point is the high complexity in the form of two electric machines, which due to their very concept have to be very powerful since the entire energy of the internal-combustion engine must be converted.

Due to the double energy conversion, the series hybrid is only fuel-efficient under certain operating conditions. Furthermore, this principle lacks the possibility of combining the power of the internal-combustion engine with that of the electric machine (boosting).

Power-split hybrid

A power-split hybrid drive generally has two electric machines, which on the one hand function as electric variators and on the other hand perform "hybrid" functions such as boosting and recuperation. For the variator function of the two electric machines some of the mechanical power is supplied to the first machine, which now functions as a generator and converts the mechanical power into electric power. The electric power is supplied via electric inverters to the second electric machine, which functions as a motor at a speed independent of the generator and

Figure 8: Structural design of a P2 parallel-hybrid transmission based on a conventional automatic planetary transmission for a standard drive
1 From driving engine,
2 Torsion damper,
3 Electric machine or generator,
4 Multi-disk brakes,
5 Planetary-gear sets,
6 Multi-disk clutches,
7 To final-drive unit.

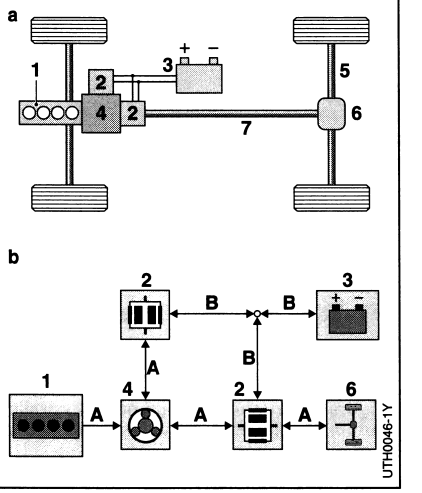


supplies its power to the drivetrain again in mechanical form (Figure 9).

The power split (or power summation) is effected by a planetary-gear set. For the input-power split the internal-combustion engine is connected to an element of the planetary-gear set (e.g. the carrier), the generator to a second element (e.g. the sun), and the second electric machine to the output and the third element (e.g. the ring gear). For the output-power split the internal-combustion engine and the generator engage an element of the planetary-gear set and the second electric machine engages a second element of the planetary-gear set, while the output is connected to the third element of the planetary-gear set. The electric component in the power-split hybrid is varied by activation of the electric machine. This provides for continuously variable starting and continuously variable ratio adjustment. The battery can be charged or discharged by varying the generated and supplied electrical energy of the two electric machines.

Standing against the high driving comfort of this concept is the high complexity in the form of two electric machines, which due to their very concept have to be very powerful as they must convert a large proportion of the energy. There is no provision for modularity with an existing basic transmission without electric machines.

Figure 9: Power-split hybrid architecture using the example of input-power split
For the output-power split the second electric machine is installed not between the planetary transmission and the final-drive differential, but between the internal-combustion engine and the planetary transmission.
a) Schematic diagram of drivetrain,
b) Energy flow.
The arrows indicate the direction of the energy flow.
1 IC engine, 2 Electric machines,
3 Traction battery, 4 Planetary transmission,
5 Drive shaft, 6 Final-drive differential,
7 Propeller shaft.
A Transmission of kinetic energy,
B Transmission of electrical energy.



References
[1] Directive 2007/46/EC of the European Parliament and of the Council of 5 September 2007 establishing a framework for the approval of motor vehicles and their trailers, and of systems, components and separate technical units for such vehicles (Framework Directive) (text of importance to the EEA).
[2] K. Reif; K.E. Noreikat; K. Borgeest: Kraftfahrzeug-Hybridantriebe – Grundlagen, Komponenten, Systeme. 1st Ed., Vieweg+Teubner Verlag, 2012.

Drivetrain with electric drives

In a vehicle with a purely electric drive, one or more electric motors (e.g. one motor per vehicle axle) convert the electrical energy, which is for the most part stored in a battery, into kinetic energy to drive the vehicle. Electric motors are quieter than internal-combustion engines, are virtually vibration-free, emit no local pollutants, and have a very high efficiency. For the same maximum power output, electric motors are much more compact than internal-combustion engines.

Pure battery-operated electric drive

An electric motor develops its maximum torque at a standstill and has a very wide effective rotational-speed range. These attributes mean that an electric vehicle can in most cases manage with a fixed ratio between the electric driving motor and the driven wheels (see Figure 1 and Figure 2). The traction battery of electric drives is based for the most part on lithium-ion technology with voltage levels of 300 to 800 V. Since the energy density of batteries is significantly lower than that of liquid fuel, in a purely electric drive either the ranges are limited on one battery charge or the batteries are much bigger and heavier than a full fuel tank on a vehicle with a conventional drive.

Figure 1: Structural design of an electric final-drive unit

1 Power electronics for activating the electric machine,
2 Electric machine,
3 Differential.

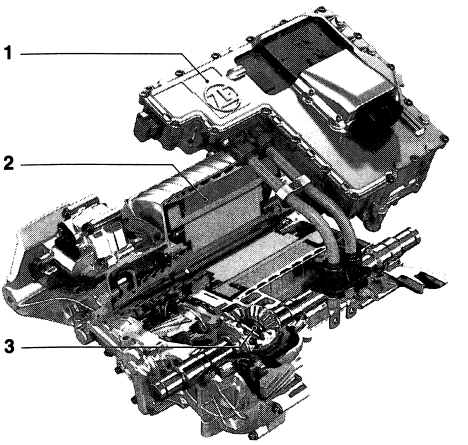


Figure 2: Schematic diagram of a vehicle with an electric drive
1 Electric machine,
2 Step-up gear unit (fixed ratio),
3 Final-drive differential, 4 Drive shaft,
5 Traction battery.

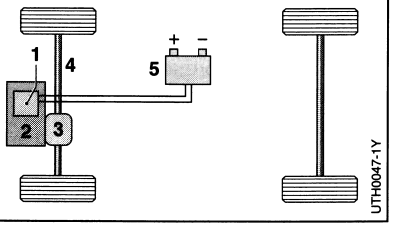


Figure 3: Schematic diagram of an electric drive with range extender in a parallel arrangement
1 Range extender, 2 Electric machine,
3 Manual transmission,
4 Final-drive differential,
5 Drive shaft, 6 Traction battery.

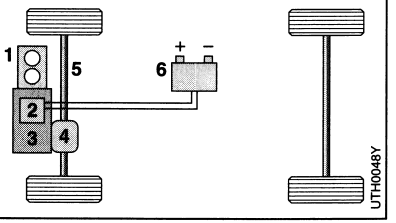
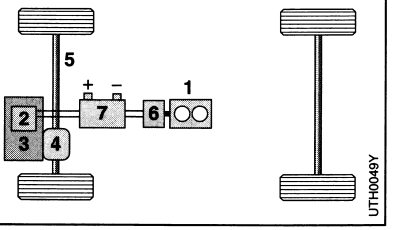


Figure 4: Schematic diagram of an electric drive with range extender in a series arrangement
1 Range extender, 2 Electric machine,
3 Step-up gear unit (fixed ratio),
4 Final-drive differential, 5 Drive shaft,
6 Generator, 7 Traction battery.



Electric drive with range extender

This form of drive corresponds to a parallel or a series hybrid with the difference that the range-extender unit of internal-combustion engine and if necessary existing generator is not as powerful as the electric driving motor, but only delivers the power required on average for limited driving operation or for the recharging process. The emphasis is on electric driving and the range-extender unit only enters into operation when the battery is almost flat. When the battery is flat, the range extender enables the vehicle to continue to be driven as if it were running on liquid fuel. Here, with a parallel arrangement (see Figure 3), the range-extender unit ensures propulsion by means of a manual transmission with a number of gears. With a series arrangement (see Figure 4), the battery is recharged by the range extender by means of the generator and propulsion is achieved as in a purely electric drive in most cases with a fixed ratio between the electric driving motor and the drive wheels.

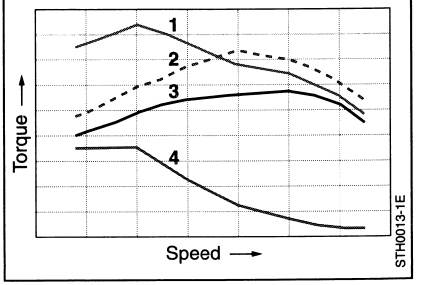
Hybrid drives

Features

A hybrid electric vehicle (HEV) uses both an internal-combustion engine and at least one electric machine for its means of propulsion. There are in this respect a multitude of drive structures which partly pursue different optimization objectives and which utilize to differing extents electrical energy to drive the vehicle (see Drivetrain with hybrid drives). Essentially three objectives are pursued with the use of electric hybrid drives:
- reduced fuel consumption and thus CO_2 emissions,
- reduced pollutant emissions, and
- increased torque and power output to improve operative dynamics.

Hybrid vehicles require an electric energy accumulator to supply the electric drive. Current solutions use a nickel-metal-hydride or lithium-ion traction battery at a comparatively high voltage level in the range of up to 800 V.

Figure 1: Torque curve of different vehicle drives
1 Hybrid drive, consisting of 3 and 4,
2 Standard IC engine with 1.6 *l* displacement,
3 Turbocharged IC engine with 1.2 *l* displacement,
4 Electric machine, 15 kW.



Advantages
The electric drive consists of an electric machine and a pulse-controlled inverter (see Electric machines for motor-vehicle propulsion). The electric drive offers constantly high torques at low rotational speeds. In this way, it ideally supplements the internal-combustion engine, whose torque only starts to increase at mid-range rotational speeds. The electric drive and internal-combustion engine together are thus able to deliver a high dynamic response from every driving situation (Figure 1).

The combination of electric and combustion-engine drives has the following advantages over a conventional drivetrain:
- The assistance provided by the electric drive makes it possible to operate the internal-combustion engine predominantly in the range of its best efficiency or in ranges in which only low pollutant emissions occur (operating-point optimization occurs).
- The combination with an electric drive facilitates the use of a smaller internal-combustion engine while retaining the same overall power output (power-neutral downsizing).
- Furthermore, a higher-geared transmission can be used while retaining the same levels of driving performance. This shifts the combustion engine's operating points into ranges with improved efficiency (downspeeding).
- Through generator operation of the electric machine, it is possible when braking to convert part of the vehicle's kinetic energy into electrical energy (recuperation). The electrical energy is stored in the energy accumulator and can be used at a later stage to drive the vehicle.
- In certain drive structures the electric drive can be use for purely electric driving. Here the internal-combustion engine is switched off and the vehicle is locally operated with zero emissions.

Functions
Depending on the operating state and required drive power, the internal-combustion engine and the electric drive contribute to propelling the vehicle to different extents. The hybrid control system determines the distribution of power between the two drives. The way in which the internal-combustion engine, the electric drive and the energy accumulator (battery) interact determines the different functions (see Drivetrain with hybrid drives).
- Start/stop function: Switch-off of the internal-combustion engine when the vehicle is at a standstill.
- Recuperation: Utilization of the generator braking torque when decelerating – the electric machine converts kinetic energy into electrical energy, which is fed to the battery.
- Hybrid driving: Internal-combustion engine and electric machine jointly deliver the desired drive power. In motor mode the electric machine supports the internal-combustion engine (boosting), while the battery is discharged. In generator mode the internal-combustion engine delivers a higher power so that the battery is charged.
- Purely electric driving: The vehicle is driven solely by the electric machine, while the internal-combustion engine is switched off.
- Charging from the mains-supply network: The battery not only is charged in generator mode, but can also be charged for example overnight from the mains-supply network.

Classification based on proportion of electric power
Hybrid vehicles can be divided into different categories based on the effected functions (see Drivetrain with hybrid drives):
- start/stop system,
- mild hybrid,
- full hybrid,
- plug-in hybrid.

Drive structures

There are different ways of configuring the internal-combustion engine, transmission and electric machines in hybrid vehicles. The different drive structures can be divided using possible energy flows into three categories (see also Drivetrain with hybrid drives).

Parallel hybrid drive
In parallel hybrid drives the internal-combustion engine and an electric machine contribute to driving the vehicle independently of one another. The two energy flows from the internal-combustion engine and the battery therefore run parallel to each other, the two individual components of power adding up to a total drive power. Parallel hybrid drives come in a mild-hybrid variant (with start/stop function, regenerative braking and hybrid driving) or in a full-hybrid variant (additionally with electric driving).

A fundamental advantage of the parallel hybrid is the possibility of maintaining the conventional drivetrain in wide ranges. The extent of development and installation work for parallel drive structures is lower when compared with series and power-split structures, in that for the most part only one electric machine with lower electric power is required and the adaptations necessary in converting a conventional drivetrain are fewer.

Parallel hybrid with one clutch
In the variant shown in Figure 2 the electric machine is connected directly to the internal-combustion engine. In contrast to series and power-split drive structures, the speed of the internal-combustion engine cannot be adjusted independently of the speed of the electric machine. In vehicle deceleration phases the internal-combustion engine cannot be decoupled from the electric machine and is therefore always engaged. The associated drag torque of the engine reduces the potential for regenerative braking (recuperation).

Purely electric driving is not possible with this drive structure. The electric drive can in fact be used as the sole drive source, but the internal-combustion engine is always engaged when driving. The

electric drive can be used to assist the internal-combustion engine and thereby improve dynamic performance.

Parallel hybrid with two clutches
A parallel full hybrid can be structured in several ways. The obvious structure is the following extension (Figure 3): integrated between the internal-combustion engine and the electric machine is an additional clutch which allows the engine to be cut in and out as required. Purely electric driving is therefore possible. The internal-combustion engine can also be decoupled in deceleration phases. Firstly, it increases the potential for regenerative braking. Secondly, it permits sailing operation, whereby the vehicle rolls freely and is slowed down only by aerodynamic drag and rolling friction.

To make this drive structure acceptable, it is very important to enable the

Figure 2: Parallel hybrid with one clutch
a) Drive structure,
b) Energy flow.
1 IC engine,
2 Fuel tank,
3 Electric machine,
4 Clutch,
5 Transmission,
6 Final-drive unit,
7 Pulse-controlled inverter,
8 Battery.

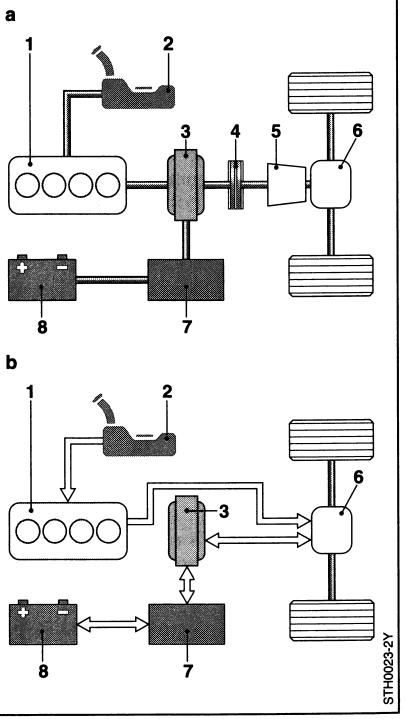
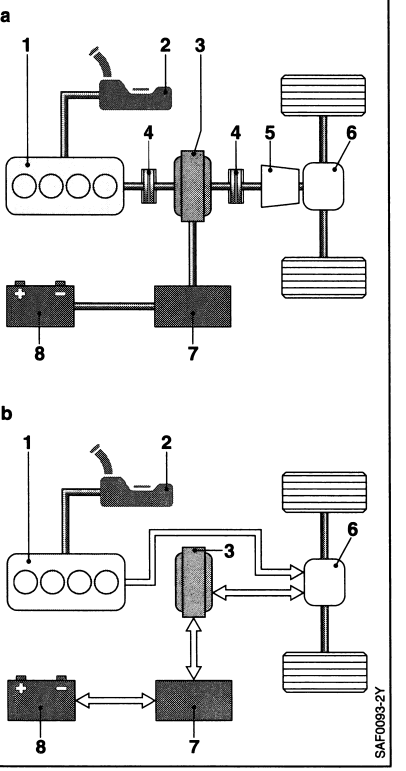
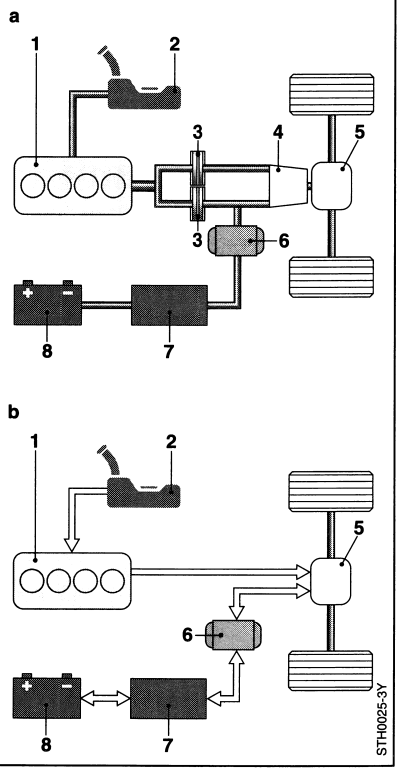


Figure 3: Parallel hybrid with two clutches
a) Drive structure,
b) Energy flow.
1 IC engine,
2 Fuel tank,
3 Electric machine,
4 Clutch,
5 Transmission,
6 Final-drive unit,
7 Pulse-controlled inverter,
8 Battery.



internal-combustion engine to be started from electric driving without compromising comfort. There are two different ways of achieving this. In the first option the internal-combustion engine is started when the interrupting clutch is open by a separate starter and there is no unwanted effect on vehicle motion. This however calls for a separate starter, which can actually be dispensed with in a hybrid vehicle. Another option is to activate the internal-combustion engine, the electric drive and the clutch in such a way that the effect on vehicle motion is compensated for during engine starting. To this end an intelligent control system requires access to measured values from the internal-combustion engine, the electric drive and the clutch. The clutch must be able to adapt automatically to the changing conditions in ongoing operation and to follow the control-system inputs.

Figure 4: Parallel hybrid with dual-clutch transmission
a) Drive structure,
b) Energy flow.
1 IC engine,
2 Fuel tank,
3 Clutch,
4 Dual-clutch transmission,
5 Final-drive unit,
6 Electric machine,
7 Pulse-controlled inverter,
8 Battery.



Parallel hybrid with dual-clutch transmission
Installing the additional clutch between the internal-combustion engine and the electric machine results in the drivetrain being increased in length. In some vehicles the installation space required for this drive configuration is not available. This problem can be remedied by integrating the electric machine in a dual-clutch transmission (Figure 4). The electric machine is no longer connected to the crankshaft of the internal-combustion engine, but is instead connected to a subunit of the dual-clutch transmission. In this configuration there is no additional clutch between the engine and the electric machine. Purely electric driving with the engine stopped is possible by opening the double clutch of the transmission. Therefore this configuration is also a parallel full hybrid. Depending on the gear selected in the transmission subunit, with the electric machine a different transmission ratio is possible between the engine and the electric machine. This produces an additional degree of freedom for the hybrid control, which can be used to further reduce fuel consumption.

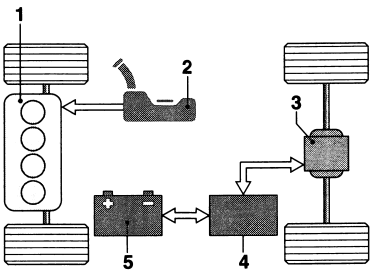
Axle-split parallel hybrid
Another parallel drive structure is produced by electrifying a separate axle (Figure 5). Here, a conventional drivetrain comprising internal-combustion engine and transmission on a powered axle is combined with an electrically powered axle. The drive configuration becomes a full hybrid as soon as the internal-combustion engine can be switched off and decoupled while the electric drive propels the vehicle. A semi-automatic transmission and a start/stop system for the internal-combustion engine are required

for this purpose. This drive structure is classed as a parallel hybrid drive because the individual components of power from the engine and the electric drive add up. In contrast to the drive structures described above, the addition point is not within the drivetrain, but on the plane of the powered wheels.

In this case the traction battery is recharged by regenerative braking. The traction battery cannot be recharged when the vehicle is stationary.

The interaction of the internal-combustion engine and the electric drive can deliver an all-wheel drive to the vehicle. Distribution of the drive torques can be adjusted by specific activation of the electric drive within wide limits. Permanent all-wheel drive can however only be realized when not only the original electric drive is supplied by the battery, but also a second electric machine can deliver the required electrical energy. This second electric machine directly connected to the internal-combustion engine (either on the crankshaft or in the belt drive) means not only that permanent all-wheel drive is realized but also that the battery is recharged even when the vehicle is stationary.

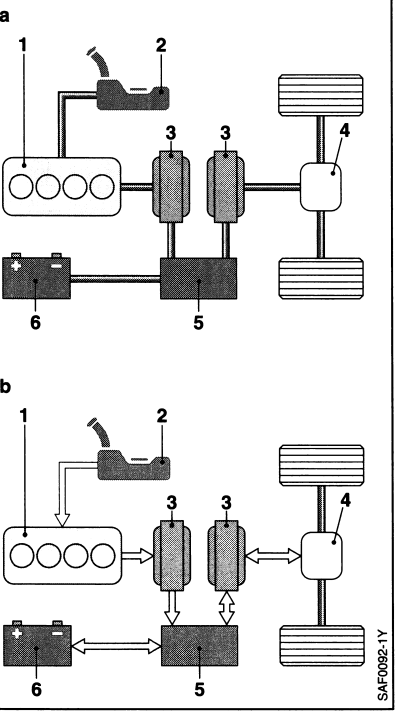
Figure 5: Electrification of a separate axle (axle-split parallel hybrid)
1 IC engine,
2 Fuel tank,
3 Electric machine,
4 Pulse-controlled inverter,
5 Battery.



Series hybrid drive

In series hybrid vehicles (Figure 6) the internal-combustion engine drives an electric machine, which operates as a generator. The electric power generated in this way is made available together with the battery power to a second electric machine, which drives the vehicle. From the perspective of the energy flows, a series connection exists in this case. A series hybrid is always a full hybrid, since all the required functions (start/stop function, regenerative braking, hybrid driving, electric driving) are possible.

Figure 6: Series hybrid drive
a) Drive structure,
b) Energy flow.
1 IC engine,
2 Fuel tank,
3 Electric machine,
4 Final-drive unit,
5 Pulse-controlled inverter,
6 Battery.



In view of the fact that there is no mechanical connection between the internal-combustion engine and the powered wheels in the series hybrid, this drive structure offers some advantages. Thus, a conventional range transmission is not required in the drivetrain. This creates free space for packaging the overall drive. In addition, starting the internal-combustion engine from electric driving does not cause any unwanted effect on vehicle motion. The chief advantage when driving is the ability to freely select the operating point of the engine. This supports fuel-saving and low-emission operational management of the vehicle. The engine can also be optimized to a limited operating range.

A downside of the series hybrid is that electrical energy has to be converted twice. The losses incurred through double energy conversion are higher than in the case of a purely mechanical transfer by a transmission. Furthermore, two electric machines of the same power category as the internal-combustion engine are needed to transmit the engine's power.

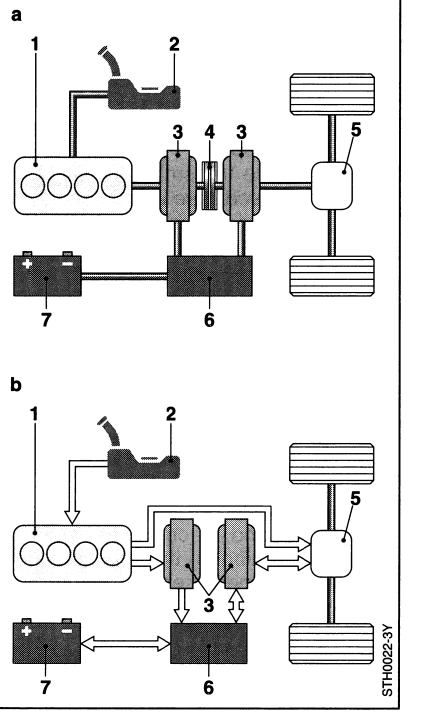
At low speeds a series hybrid, despite the higher losses, offers a fuel-consumption advantage because here the positives afforded by the ability to freely select the engine's operating point outweigh the negatives. The higher losses are predominant at medium and higher speeds.

Series hybrids are currently used predominantly in diesel-electric locomotives and city buses. In the passenger-car sector series drive structures are being encountered more and more frequently in electric vehicles whose ranges are extended where necessary by an internal-combustion engine as a "range extender".

Series-parallel hybrid

A series hybrid is extended to become a series-parallel hybrid (Figure 7) by establishing a mechanical connection between the two electric machines which is connected or separated by a clutch. The series-parallel hybrid can utilize the advantages of the series hybrid at low speeds and circumvent the disadvantages at higher speeds by closing the clutch. When the clutch is closed, the series-parallel hybrid behaves like a parallel hybrid. Since the double energy conversion is limited to the range of lower speeds and power

Figure 7: Series-parallel hybrid drive
a) Drive structure,
b) Energy flow.
1 IC engine,
2 Fuel tank,
3 Electric machine,
4 Clutch,
5 Final-drive unit,
6 Pulse-controlled inverter,
7 Battery.



outputs, smaller electric machines are sufficient for the series-parallel hybrid than are required for the series hybrid. Compared with the series hybrid, the advantage in packaging is lost on account of the mechanical connection between the internal-combustion engine and the powered wheels. Compared with the parallel hybrid, two electric machines are required for the same task.

Power-split hybrid drive

Power-split hybrid vehicles combine features of parallel and series hybrid vehicles with those of a power split. Some of the engine power is converted by a first electric machine into electric power, the remainder – together with a second electric machine – driving the vehicle. A power-split hybrid is always a full hybrid, since all the required functions (start/stop function, regenerative braking, hybrid driving, electric driving) are possible.

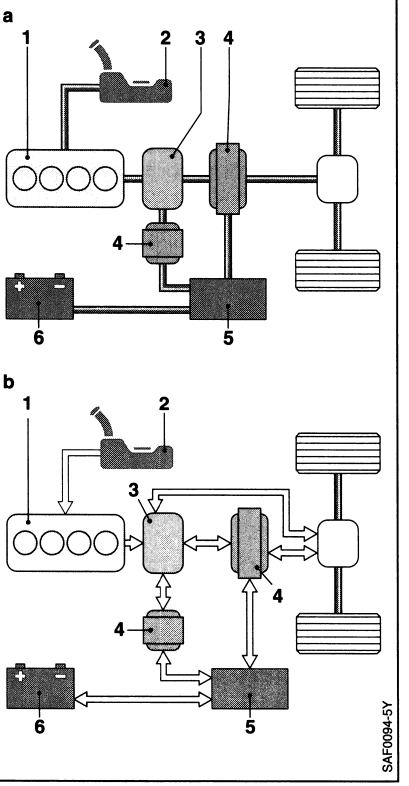
The design is shown in Figure 8. The central element is a planetary-gear set, which is connected via its three shafts to the internal-combustion engine and two electric machines. Because of the kinematic boundary conditions at the planetary-gear set, the engine speed can be adjusted within certain limits independently of the driving speed. Following a continuously variable transmission (CVT), the terminology used here is an electric continuously variable transmission (ECVT).

Some of the engine power is transmitted by the planetary-gear set via a mechanical path to the powered wheels. The remainder of the power is transmitted via an electrical path with double energy conversion to the powered wheels. Similarly to the series hybrid, the electric transmission path can be used in the case of low power demands. For higher power demands, the mechanical transmission path is also available. However, it is not possible to switch arbitrarily between the mechanical and electrical transmission paths. Depending on the configuration of the planetary-gear set, the electric machine and the internal-combustion engine, without additional transmissions only certain combinations between the mechanical and electrical transmission paths are ever possible. In this way, the power-split hybrid can achieve substantial fuel savings at low and medium speeds. No additional fuel savings can be achieved at high speeds.

Similar to the series hybrid, the power-split hybrid requires electric machines with relatively high power outputs in the range of the installed engine power.

The power-split hybrid can be expand by mechanical, fixed gears by using a second planetary-gear set. The mechanical

Figure 8: Power-split hybrid drive
a) Drive structure,
b) Energy flow.
1 IC engine,
2 Fuel tank,
3 Planetary transmission,
4 Electric machine,
5 Pulse-controlled inverter,
6 Battery.

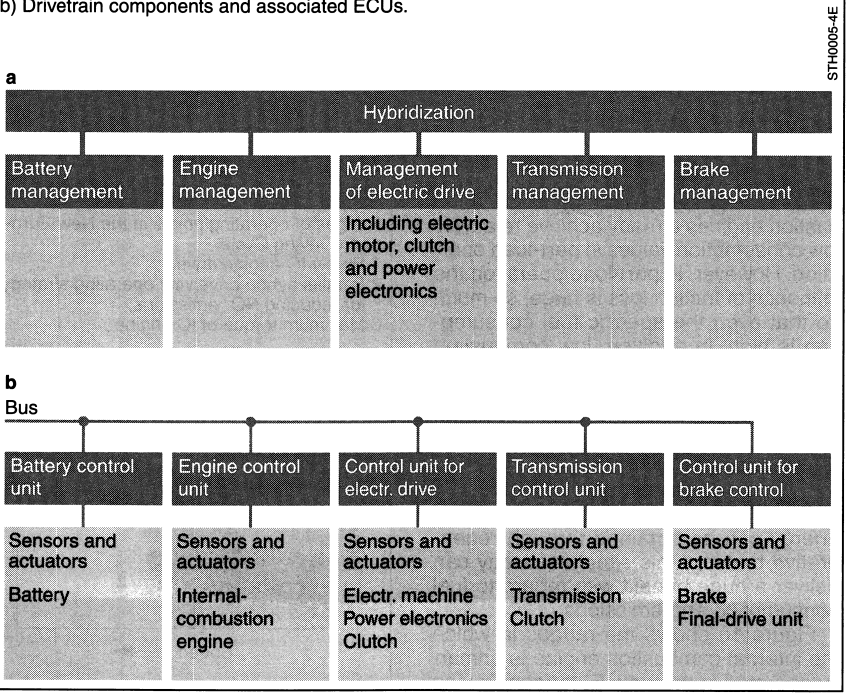


complexity increases, while the electrical complexity decreases. Smaller electric machines are then sufficient for a comparable concept. Furthermore, fuel consumption at medium and higher speeds can be improved.

Control of hybrid vehicles

The efficiency which can be achieved with the relevant hybrid drive is crucially dependent on the higher-level hybrid control. Figure 9 uses the example of a vehicle with a parallel hybrid drive to show the function and software structure and the networking of the individual components and ECUs in the drivetrain. The higher-level hybrid control coordinates the entire system, the subsystems of which have their own control functions. These are battery management, engine management, management of the electric drive, transmission management, and management of the brake system.

Figure 9: Networking of control systems in the drivetrain
a) Function and software structure,
b) Drivetrain components and associated ECUs.



In addition to pure control of the subsystems, the hybrid control also includes an operating strategy which optimizes the way in which the drivetrain is operated. The operating strategy brings influence to bear on the consumption- and emission-reducing functions of the hybrid vehicle, i.e. on start-stop operation of the engine, regenerative braking, and hybrid and electric driving.

Operating strategies for hybrid vehicles

The operating strategy determines how the drive power is shared between the engine and the electric drive. It decides the extent to which the potentials for fuel saving and for reducing the emissions of a vehicle are utilized. The operating strategy must also implement the different hybrid functions such as regenerative braking and hybrid and electric driving.

Selection of and switching between the individual states is performed in accordance with numerous conditions relating to, for example, the accelerator-pedal position, the state of charge of the battery, and the vehicle's current speed. The components in the hybrid vehicle behave differently, depending on the optimization objective (e.g. fuel saving or emission reduction).

Operating strategy for reducing NO_x

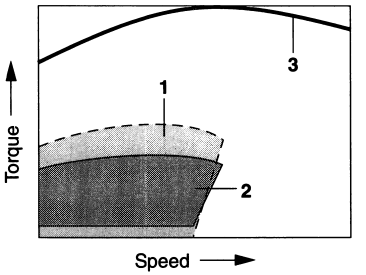
Vehicles with lean-running internal-combustion engines already achieve relatively low consumption values in part-load operation. However, in part-load operation the influence of friction loss is huge, so much so that even the specific fuel consumption is high. In addition, low combustion temperatures and a local oxygen deficiency in the low part-load range result in high carbon-monoxide and hydrocarbon emissions. Already an electric drive with relatively low power can replace the internal-combustion engine as drive in the low load range. If the required electrical energy can be obtained through regenerative braking, this simple strategy can deliver a huge benefit with regard to fuel consumption and emissions.

Figure 10 shows the ranges in which the internal-combustion engine is primarily operated in the New European Driving Cycle (NEDC). A passenger-car diesel engine is operated both at low part load (i.e. with poor efficiency levels and high HC and CO emissions) and at medium/higher load (i.e. in the range of high NO_x emissions).

Figure 10 shows furthermore the range of operating points for a parallel hybrid which bypasses low internal-combustion engine loads through purely electric driving or through load-point increase. This on the one hand reduces the fuel consumption, but on the other hand reduces the CO, HC and NO_x emissions – which are high in this range. To achieve a further lowering of NO_x emissions, it is possible to lower the load points in the medium load range by simultaneously operating the electric drive and the engine.

Figure 10: Operating strategy for reducing NO_x emissions

Ranges of operating points in the New European Driving Cycle.
1 Purely IC-engine drive,
2 Parallel hybrid drive with operating strategy for reducing NO_x emissions,
3 Maximum torque of IC engine.



Operating strategy for reducing CO_2

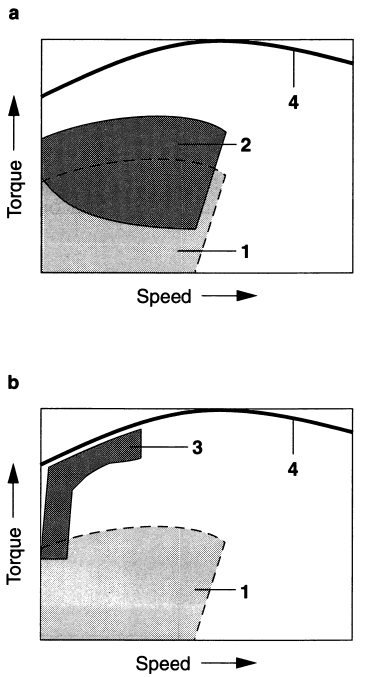
With vehicles with stoichiometrically running gasoline engines, extremely low emission values can be realized on account of the three-way catalytic converter used. In these vehicles the focus is on reducing the fuel consumption and thus also the CO_2 emissions. Figure 11 shows for different drive structures a possible optimization of the engine's operating range with regard to minimal CO_2 emissions.

In the New European Driving Cycle (NEDC) internal-combustion engines in conventional vehicles are operated at low part load and thus with poor efficiency. In vehicles with parallel hybrid drives low engine loads can be avoided by means of purely electric driving (Figure 11a). Since the required electrical energy as a rule cannot be obtained exclusively through recuperation, the electric machine is then operated as a generator. This results, when compared with a conventional vehicle, in a shift in engine operation to higher loads and thus in better levels of efficiency.

In the case of the power-split hybrid vehicle (Figure 11b), the operating range of the internal-combustion engine is subject to greater limitations than the parallel hybrid vehicle. As a rule it is operated as a function of engine speed at the load at which the entire drivetrain operates under optimum energy conditions.

Figure 11: Operating strategy for reducing CO_2 emissions

Ranges of operating points in the New European Driving Cycle.
a) Comparison of a purely IC-engine drive with a parallel hybrid drive,
b) Comparison of a purely IC-engine drive with a power-split drive.
1 Purely IC-engine drive,
2 Parallel hybrid drive,
3 Power-split hybrid drive,
4 Maximum torque of IC engine.



Regenerative braking system

In the case of regenerative braking (also known as recuperation), the vehicle's kinetic energy during deceleration is converted by the electric machine, which is operated as a generator for this purpose, into electrical energy. In this way, some of the energy which is normally lost as frictional heat during braking is fed in the form of electrical energy to the battery and then utilized.

Drag-torque simulation
A simple way of realizing regenerative braking is drag-torque simulation. Here the electric machine is operated as a generator as soon as the driver releases the accelerator. The driver does not need to press the brake pedal for this purpose. In the case of a full hybrid, the internal-combustion engine is decoupled and the electric machine exerts a generator torque of the magnitude of the engine's drag torque.

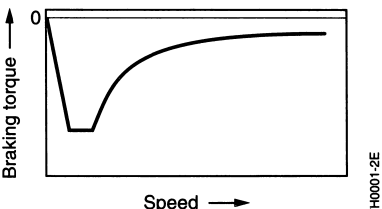
If the engine cannot be decoupled (as in the case of a mild hybrid), alternatively a lower generator torque can be exerted on the drivetrain in addition to the engine drag torque (drag-torque increase). The vehicle behavior changes only slightly compared with a non-hybrid vehicle.

Simple regenerative braking system
During braking the electric machine can exert an additional generator torque as well as drag-torque simulation and increase. The vehicle decelerates with the brake pedal in the same position faster than a comparable conventional vehicle. The available generator torque depends on the driving speed, the selected gear and the state of battery charge. Therefore the vehicle can be subject to different levels of braking response, even with the brake pedal in the same position. This difference in braking response is perceived by the driver to be all the more disruptive, the greater the generator torque is part of the vehicle deceleration. For this reason, only small power levels can be recuperated with this simple regenerative braking system.

Cooperatively regenerative braking system
The service-brake system must be modified in the case of higher decelerations in order to further exploit the kinetic energy. To this end all or some of the service-brake friction torque must be replaced with a regenerative braking torque without the vehicle deceleration changing when the brake-pedal position and brake-pedal force are kept constant. This is realized in a cooperatively regenerative braking system, where the vehicle control and the braking system interact in such a way that constantly as much friction braking torque is recovered as generator braking torque can be replaced by the electric machine.

During a braking operation the maximum generator torque that can be achieved by the electric motor changes continuously over a wide speed range (Figure 12). This is due to the fact that the electric-motor power (i.e. torque times speed) is constant in this range. Only at low speeds is there a range of constant maximum electric-motor torque. If the speed continues to drop, the regenerative braking torque that can be achieved again drops to zero. A constant torque is required at the wheels for a constant deceleration of the vehicle. If the electric motor operated as a generator is exploited to its limit torque, the friction braking torque must be continuously reduced with decreasing driving speed because

Figure 12: Maximum torque path for electric-machine generator mode



the generator torque increases. Figure 13 shows by way of example a braking operation of a power-split hybrid. At the start of the maneuver the generator torque is increased until it reaches its maximum level (if this corresponds to the demanded total braking torque). Towards the end of the braking operation the generator torque is reduced and replaced completely by a friction braking torque because the electric motor can no longer deliver a generator torque at very low rotational speeds (Figure 12). During the braking operation from high driving speed to a standstill, therefore, the distribution between friction braking torque and regenerative braking torque is continuously adapted with constant pedal actuation.

Braking-force distribution

As with the design of conventional brake systems, so too with the design of a regenerative brake system the braking-force distribution between front and rear axles is of crucial importance to the vehicle's directional stability. As deceleration increases, the normal force on the front wheels increases while the normal force on the rear wheels decreases. If the electric motor is connected to the front wheels, a greater wheel torque can also be transmitted with increasing deceleration and thus increasing normal force on the front axle. Therefore, to achieve directional stability, friction-coefficient utilization of the front axle should not exceed that of the rear axle. On a rear-wheel-driven vehicle or a vehicle with an electrified rear axle (i.e.. internal-combustion engine on the front axle and electric motor on the rear axle) the transmittable regenerative torque decreases with increasing deceleration.

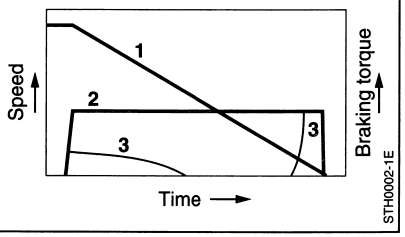
The electric power of the battery also has a major influence on the utilization of recuperation since it is the limiting element for absorbing electrical energy from the vehicle movement. The maximum possible purely recuperative deceleration increases with increasing energy-accumulator power. For vehicle-stability reasons, however, a large generator braking torque can only be transmitted on the front axle.

Influence on driving-dynamics control

Since the regenerative brake system influences braking stability, the driving-dynamics control must be adapted to the changed physics of driving. It is expedient to suppress the regenerative component of braking when unstable driving states or excessively high brake slip are detected and to realize deceleration and stabilization interventions solely by means of the friction-braking system. Otherwise instabilities and drivetrain vibrations could disrupt optimum wheel-slip control.

Figure 13: Cooperative braking maneuver for a power-split hybrid

1 Driving speed,
2 Total braking torque,
3 Friction braking torque.



Boost-recuperation system 48 V

Motivation

The constantly increasing demands to optimize the CO_2 emissions of current vehicles present the vehicle manufacturers with huge challenges. Aside from friction and combustion-process optimization, electrification of the drivetrain is an effective measure for reducing CO_2 emissions. The 48 V boost-recuperation system (BRS) from Bosch offers a cost-optimized electrification approach in this connection.

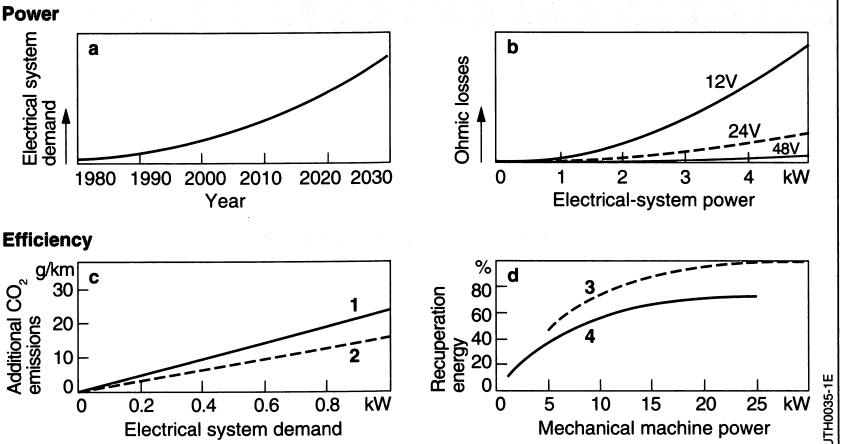
Electrical power demand

The electrical power demand in passenger cars is rising continuously and will continue to increase in the next few years due to the use of an increasing number of powerful electrical comfort components (Figure 1a). The electrical energy required to supply the electrical loads/consumers is currently generated in conventional vehicles by a 14 V alternator which is connected via a belt to the internal-combustion engine's crankshaft. The load on the internal-combustion engine for generating current increases the fuel consumption. For a mid-size vehicle the fuel consumption – and hence the CO_2 emissions – for example in the New European Driving Cycle (NEDC) at a electrical-system load of 500 W increases by approximately 0.5 *l* per 100 km (Figure 1c).

However, the automotive industry is aiming to achieve a significant reduction in fuel consumption and consequently CO_2 emissions. One way of reducing the CO_2 emissions on the engine side is to switch off the internal-combustion engine when it is not needed to propel the vehicle. The electrical energy required to supply the vehicle electrical system during this time is delivered by the battery. Some functionalities can consequently also be realized

Figure 1: Development of electrical-system demand and its influence on CO_2 emissions

a) Increase in electrical-system power (trend),
b) Reduction of losses (trend),
c) Influence on CO_2 emissions,
d) Share of recuperation energy.
1 New European Driving Cycle (NEDC),
2 Worldwide Harmonized Light Vehicles Test Cycle (WLTC),
3 Not including drag losses of IC engine,
4 Including drag losses of IC engine.



with the existing 12 V electrical-system technology and its components, for example by start/stop operation when the vehicle is at a standstill or even by start/stop sailing (coasting) with manageable technical outlay. These functionalities are in part already state of the art.

Plug-in hybrids approach

Furthermore, recent years have witnessed the development of plug-in hybrids with whose electric drives and electric high-power accumulators the vehicles can drive for a number of kilometers by purely electric means. Moreover, the large electric power output facilitates engine boost assistance and during braking a recuperation of the vehicle's kinetic energy into electrical energy. However, the process of satisfying the technical requirements imposed on a plug-in hybrid is very cost-intensive due to the required vehicle system voltages of several hundred volts, and up to now it has not been possible to carry this through to any significant effect on the market in the volume segment.

Electrification requirements

The purpose of the following deliberation is to consider what requirements vehicle electrification should satisfy in order on the one hand to remain cost-effective and on the other hand already to cover some of the functionalities of a plug-in hybrid. Considerations shall in this connection focus on fuel economy by means of mild hybridization. Figure 1 shows two core requirements for this purpose.

Reduction of electrical losses

The first core requirement is reducing the electrical losses in spite of increasing electrical-system load. This can be achieved by increasing the voltage in the vehicle electrical system. Increasing the voltage enables the current to be reduced with the same power output and thereby reduce the electrical losses (Figure 1b). Up to a voltage limit of 60 V (direct voltage) the voltage network in vehicle manufacturing is classed as low voltage and does not require any specific shock-protection measures, which provides for cost-effective implementation.

Recuperation

The second core requirement is aimed at reducing fuel consumption by converting kinetic vehicle energy into electrical energy (recuperation). With an electrical machine power output of approximately 12 kW already a large proportion (roughly 80 %) of the available kinetic energy can be recuperated and accordingly energy can be saved to generate current in the vehicle. Depending on the electrical-system load and the vehicle, even more energy can be recuperated than is needed for the electrical-system supply. In this case it is expedient to export this energy in the form of mechanical torque assistance to the internal-combustion engine back to the crankshaft.

Implementation at 48 V level

The voltage level of 48 V, as can been seen in Figure 1, on the one hand offers the required performance for the above-derived core requirements; on the other hand, it can be implemented with technically manageable outlay. Derived from this, it makes sense also to increase the generation of power in the vehicle from 12 V to 48 V. However, it is now necessary to assess the extent to which the requirements regarding a 48 V alternator change compared with a conventional 12 V alternator. Furthermore, to deliver hybrid functionalities, it is necessary to facilitate, aside from the pure generator function, motor mode as well, which calls for an electric machine with an operating voltage of 48 V. The question now arises as to the machine concept with which these new requirements can be implemented in the most efficient way.

Requirements imposed on a 48 V electric machine

The requirements which are imposed on a 48 V electric machine depend not least on the installation location and consequently on the drivetrain topology. There are basically three different mounting positions for the electric machine in the drivetrain (see Hybrid drives):

- on the internal-combustion engine with belt connection to the crankshaft (P0 or P1 topology, here predominantly P0),
- between the internal-combustion engine and the transmission inlet,
- at the transmission outlet with direct axle connection (P3 or P4 topology).

Figure 2 shows the different possible installation positions for the electric machine.

Topologies

Connection to the internal-combustion engine

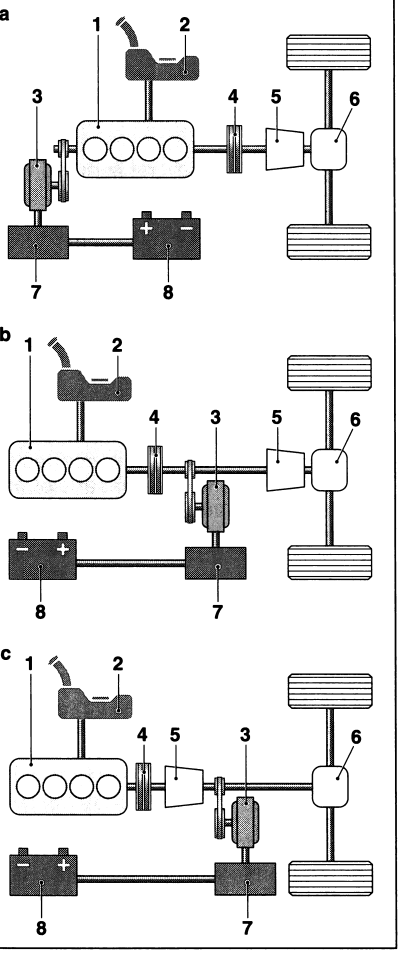
In the case of belt connection, the electric machine replaces the conventional alternator. It therefore stands to reason for the electric machine to be designed with a machine concept corresponding to the conventional alternator as a claw-pole machine. Consequently, it also makes sense to keep the gear ratio of approximately 3:1 between the electric machine and the crankshaft and thus to leave the rotational-speed spread as it is for a conventional alternator.

The electric machine can effectively support the internal-combustion engine on account of its characteristic torque curve as shown in Figure 3 during acceleration operations particularly in the lower rotational-speed range. As the speed increases, the power output of the claw-pole machine can be kept constant and consequently the entire acceleration phase uniformly supported. Thanks to the characteristic torque curve, starting the internal-combustion engine after a standstill phase is possible via the belt drive and this happens more quietly and faster than during conventional pinion starting.

The electric machine can with this topology perform the conventional alternator function even when the vehicle is at a standstill and thereby reliably supply the vehicle electrical system. On the proviso that the electric machine has a suitably compact design and no additional water

Figure 2: Installation positions of the 48 V electric machine

a) Connection to IC engine (example of a P0 topology),
b) Connection to transmission inlet,
c) Connection to transmission outlet (example of a P3 topology).
1 IC engine, 2 Fuel tank,
3 Electric machine, 4 Clutch,
5 Transmission, 6 Final-drive unit,
7 Pulse-controlled inverter, 8 Battery.



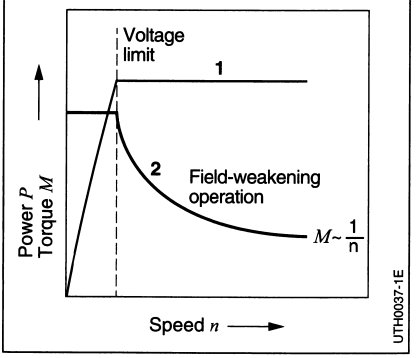
cooling is required, mounting the electric machine in the original alternator position entails only minimal additional integration outlay, since the alternator installation space can be utilized and only the belt tensioner needs to be changed.

Connection to the transmission inlet
If the electric machine is mounted directly after the clutch on the transmissions input shaft, purely electric starting and driving are possible provided the clutch remains open. The different transmission ratios are optimal for this. The rotational-speed gradients at the electric machine are however higher due to the clutch processes than with the belt connection to the internal-combustion engine.

Likewise, the additional moment of inertia acting via the transmissions input shaft must be taken into account in the design of the synchronizer rings. Reliable generator operation via the internal-combustion engine is however interrupted when the clutch is open.

Since the internal-combustion engine can be disengaged in recuperation mode, more energy can be recovered at lower vehicle decelerations than with belt connection to the internal-combustion engine. However, engine starting can only take place with the clutch closed, which hampers the transition from electric driving to conventional driving mode.

Figure 3: Torque characteristic curve of a claw-pole machine
1 Power P,
2 Torque M.



The integration outlay for this topology is high since significant changes to the transmission housing are required.

Connection to the transmission outlet
Connecting the electric machine to the transmission outlet is also feasible. However, the rotational-speed spread of the drive axle must be transmitted to the machine shaft. Electric driving can only be achieved by choosing a suitable conversion ratio between the electric machine and the drive shaft. If, at higher driving speeds, this results in the upper rotational-speed limit of the electric machine being exceeded. the electric machine must be automatic disengaged.

In principle, this topology offers the greatest efficiency during recuperation since the lowest mechanical transmission losses occur here. But this only applies provided the electric-machine power output does not decrease substantially with the speed. Otherwise the maximum possible recuperation or boost power would drop prematurely with the driving speed. If the electric machine has be to disengaged due to its rotational-speed limit at higher driving speeds, the recuperation potential disappears entirely in these high rotational-speed ranges.

The internal-combustion engine cannot be started from a vehicle standstill. As well as the additionally required conventional alternator for reliable electrical-system supply this topology must have a conventional start/stop-capable starter. The integration outlay for the electric machine must here too be gaged as very high.

Advantages
Mounting in the conventional alternator position with belt connection offers the most advantages with regard to functions and integration. It is however absolutely necessary to keep the electric machine including power electronics compact so as to avoid extensive conversion measures on the internal-combustion engine.

Choice of concept for the electric machine

Permanently excited synchronous machines and asynchronous machines which exhibit specific advantages and disadvantages depending on their use are predominantly deployed in current hybrid vehicles.

Operating strategy
The vehicle's operating strategy must be taken into account as well as the necessary power density of the electric machine. Focus is placed here on recuperation and hence on the comprehensive avoiding of continuous generator operation. This means that the required electrical system energy is covered from recuperation. Conversely, the upshot is that the electric machine used has frequent idle phases in which it is inactive. The proportion of inactive time ranges, depending on driver type and distance covered, between 5 % and 45 %, by which a power loss or drag power that is as low as possible on the part of the selected electric machine is important during these operating phases. This importance becomes even more pronounced when considering and comparing the sensitivities to CO_2 reduction and fuel consumption with regard to machine efficiency or power.

Minimizing the drag power is thus a further central requirement. Machine concepts which utilize separate excitation and can thus fully deactivate the electric machine are advantageous for this purpose.

Machine concepts
By now considering the core requirements for the machine concepts under discussion (permanently excited synchronous machine, asynchronous machine, and separately excited synchronous machine, see Electric machines), it can be recognized that synchronous machines have advantages regarding power density and their power output at high rotational speeds. A significant differentiation of the efficiencies between the machine types cannot be determined. However, as described above, it is necessary to take into account the drag losses, which turn out to be much lower with separately excited machines. Consequently, asynchronous machines, but also separately excited synchronous machines, exhibit advantages. A separately excited synchronous machine is thus best suited with regard to the requirements derived for a 48 V system.

A claw-pole machine, as is used today as a vehicle alternator, delivers this machine type precisely and also satisfies the requirements with regard to power output and installation space. Since these electric machines are produced in very large numbers, i.e. mass-produced, the additional manufacturing outlay required also remains low.

Boost-recuperation machine
The boost-recuperation machine (BRM) developed by Bosch is thus the result of a consistent derivation of the requirements imposed on a cost-effective 48 V electric machine which is based on a conventional alternator. The power of the claw-pole machine here is achieved by the voltage increase and by the use of an add-on converter which assumes the position of the rectifier of a conventional alternator. For use of the boost-recuperation machine in different vehicles and under different thermal boundary conditions the concept is scalable in design both with regard to power and torque and also with regard to the thermal rating of the electric machine.

Scalability
The specific requirements imposed on the machine torques and power outputs can vary, depending on the size and type of the internal-combustion engine and on the vehicle category. These requirement must be met. The claw-pole concept of the boost-recuperation machine permits recourse to a modular system, by means of which the electric machine can be individually adapted. A significant influencing factor with regard to the above-described requirements is the electromagnetic layout of the electric machine. This can be specifically influenced both in the stator by the number of windings and in the rotor by the number of permanent magnets used for leakage-flux compensation such that the power and torque ranges indicated in Table 1 can be achieved. It must be emphasized in particular here that the entire

range of variation can be fully utilized for the identical size. The characteristic machine data can therefore be adapted to a particular application without adapting the mounting conditions in terms of design to the internal-combustion engine. The most powerful variant furthermore exhibits a high power density which, on account of the principles involved, is significantly higher than that of other separately excited machine concepts.

The electric machine is based to a large extent on components of alternators which are mass-produced (Figure 4). The most important change as against existing alternators is the use of a converter with active switching elements (MOSFET) in place of diodes. This facilitates specific setting of the stator currents and hence operation as a motor or an alternator.

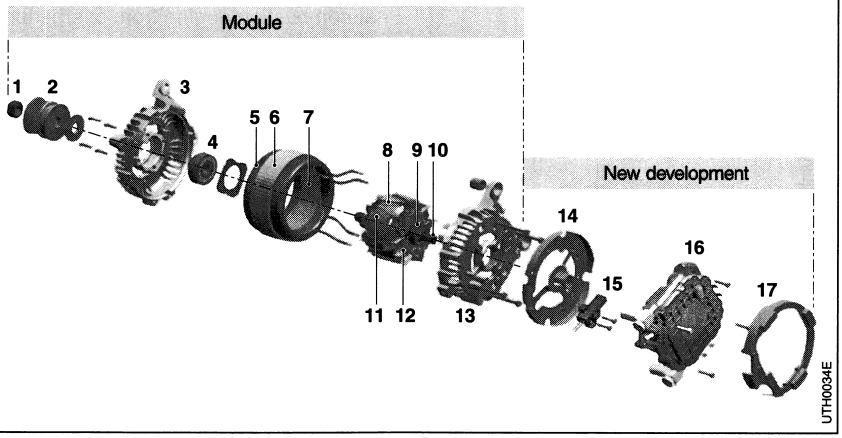
Table 1: Power and torque ranges

Starting torques	at 100 rpm	40 to 60 Nm
Maximum electric power	at 6,000 rpm	8 to 12 kW
Maximum mechanical power	at 6,000 rpm	8 to 10 kW
Electric continuous output	at idle	1.5 to 3.2 kW

Moreover, the opportunity is created on the electric-machine housing to mount a belt tensioner, which ensures the correct belt tension irrespective of the direction of rotation and operating mode. This provides for multi-quadrant operation of the boost-recuperation machine in the belt drive, i.e. operation as a motor or an alternator in both directions of rotation.

As on a conventional alternator, on the boost-recuperation machine the fan is mounted directly on the rotor. The rotation of the rotor supplies the internal components with air, ensuring that the machine components are cooled efficiently.

Figure 4: Boost-recuperation machine based on an alternator module
1 Nut, 2 Belt pulley, 3 A end shield, 4 Ball bearing,
5 Winding, 6 Stator laminations,7 Insulation paper, 8 Magnet,
9 Ball bearing, 10 Collector rings, 11 Claw pole, 12 Fan,
13 B end shield, 14 Connecting plate, 15 Brush holder,
16 Converter (Inverter), 17 Protective cap.

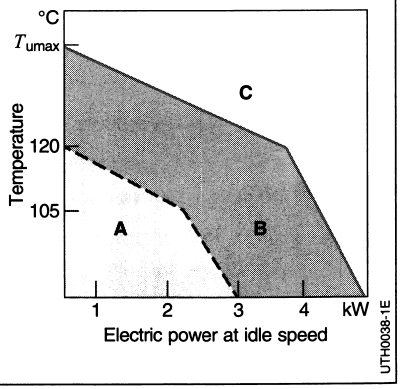


Thermal situation and rating

Air cooling

During the conversion between electrical and mechanical energy there occur in the boost-recuperation machine, on account of the principles involved, electrical, electromagnetic and mechanical losses which result in the machine components being heated. This heat must be dissipated to the surroundings. An established and cost-effective method of heat dissipation is cooling with the available ambient air in the engine compartment. Here, the air is drawn by radial fans into the boost-recuperation machine and blown out again radially via the end shield. As it passes through the boost-recuperation machine the air, on account of the temperature difference, absorbs heat from the electronics heat sink and from the electric machine.

Figure 5: Power ranges of the different cooling concepts of the boost-recuperation machine (BRM)
A BRM air-cooled,
B BRM hybrid-cooled,
C Generator continuous operation of BRM.
T_{umax} Maximum temperature.



Thermal analysis

The relevant load situations must be analyzed during use to enable the thermal situation in the boost-recuperation machine to be evaluated. During a typical driving cycle (e.g. WLTC), which contains engine starts and acceleration and braking phases, the electric-machine load varies markedly in response to the alternation of the operating modes (generator, motor, and idle) and the demanded torques. Thermal analysis shows that the temperatures of rotor, stator and electronics remain significantly below the critical temperatures. The performance of the boost-recuperation machine is thus not restricted to the relevant engine-compartment temperatures; the boost-recuperation machine can therefore be used during normal driving operation in the full power range of up to 12 kW.

Hybrid cooling concept

However, aside from typical driving operation, the boost-recuperation machine must also be evaluated with regard to thermal special cases which can arise as a result of the vehicle's engine compartment initially being strongly preheated and then the vehicle electrical system being subjected to load at idle with a high continuous output by the electrical loads/consumers. Such a situation can arise for example when during the summer months with the vehicle at a standstill the vehicle's radiator fan has to be supplied with electric power in order to prevent the internal-combustion engine from overheating.

For such extreme cases continuous outputs of in part well above 1 kW are demanded in combination with an engine-compartment temperature above 100 °C. Here the situation can arise where pure air cooling no longer suffices and the rising electronics temperature leads to down-control of the electric power. For this reason, the claw-pole concept with integrated power electronics offers as an alternative to pure air cooling the hybrid cooling concept, in which the electric machine continues to be air-cooled, but the power electronics are connected to the vehicle's water cooling circuit. With this enhanced cooling concept generator continuous outputs of up to approximately

3 kW can be achieved at idle, even at very high engine-compartment temperatures. An overview of the power ranges of the different cooling concepts is provided in Figure 5.

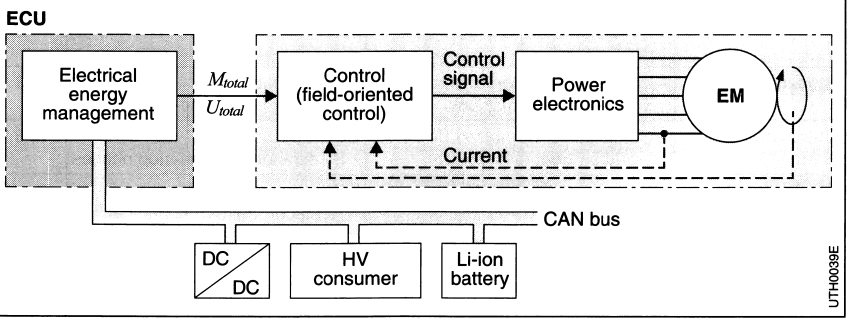
Both the air- and the water-cooled variants of the boost-recuperation machine can be realized in the same installation space thanks to the special configuration of the cooling concept. There is therefore also a scalable machine concept in the thermal respect.

Machine performance in the system network

As against an alternator, in which the electronics must rectify and regulate the voltage, the boost-recuperation machine differs by way of the additional motor mode. The individual functions of engine start and generator and motor mode are requested from the vehicle control. The choice of functions depends on the driver's requests, the driving state, the operating point of the internal-combustion engine, and the vehicle's energy management. With the control concept implemented in the boost-recuperation machine the electric machine is similarly controlled via a torque demand. It is therefore the responsibility of the vehicle control, by specifying the torque, to adapt the individual operating modes of the electric machine harmoniously to the respective driving state. This necessitates precise torque control of the electric machine, which is assumed by the microprocessor-controlled power electronics and hence switches the excitation winding and the five-phase winding accordingly. The magnetic field is set synchronously to the rotor with the aid of a rotor position sensor. The activation concept of the boost-recuperation machine in combination with the electrical 48 V system is illustrated in Figure 6.

If the boost-recuperation machine is controlled via torque input instead of via voltage, the electrical limits of the E/E system must nevertheless be observed.

Figure 6: Controlled system and activation of the boost-recuperation machine



Electrical energy management (EEM) monitors the state variables of the 48 V vehicle electrical system and the 48 V battery for this purpose. The influence of the DC/DC converter and of the 48 V loads must be taken into account here. A battery voltage varying over the state of charge (SOC) and temperature-dependent current limits necessitate a regular calculation of these variables during operation of the electric machine. The resulting current and voltage limits can be transmitted via the CAN bus to the boost-recuperation machine. If these are known to the control system, the electrical limits can be observed despite torque input. If the 48 V battery offers predicted currents and voltages for a specified time horizon, an offset against predicted torques is possible. The prevailing boundary conditions of the boost-recuperation machine are taken into account here.

Overall view of the 48 V system

The claw-pole machine with belt drive derived from the conventional alternator is suitable for mild hybridization in the belt drive and can be adapted to different vehicle requirements. The 48 V voltage level in combination with power electronics offers aside from the capacity for recuperation further properties, as shown in Table 2.

The electrical energy obtained during recuperation can be used without further voltage transformation directly to drive the electric machine. The supplementary torque thereby generated which is transmitted by belt to the crankshaft improves the vehicle's response and acceleration behavior.

The direct connection to the belt drive makes it possible, after a stop or a sail phase, to restart the internal-combustion

Table 2: Advantages and properties of the 48 V boost-recuperation system

Application	Comfort and driveability	Emission reduction	
Recuperation	– No negative repercussion	CO_2 reduction – By recuperative braking	≈6 %
Torque assistance	– Experience of driving dynamics – Improved response	CO_2 reduction – Potential for downsizing and downspeeding – Free shift-point selection in the driving cycle in a manual transmission NO_x reduction in diesel engines	≈6 %
Hybrid-like engine start	– "Engine stall protection" – Fast, quiet and low-vibration engine start	HC and CO reduction in gasoline and diesel engines	
Stop-in-gear	– Intuitive application of start/stop in manual transmissions	CO_2 reduction – By extended engine-stop phases	2 to 3 %
Start/stop Sailing	– Comfortable, no-noise driving – Vibration-free, fast restart	CO_2 reduction – By extended engine-stop phases	5 to 8 %
Electrical-system power	– Electric heating, ventilation and air conditioning – Dynamic driving-stability control	CO_2 reduction – Replacement of mechanical high-power loads/consumers with electrical loads/consumers with control to suit demand – Weight reduction	

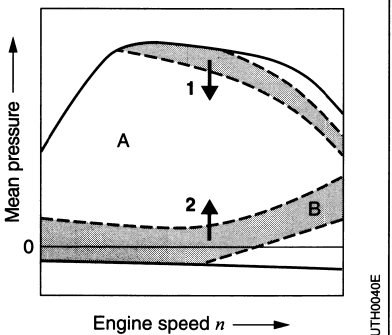
engine via the belt with very little noise and vibration and thus to increase comfort significantly. Even stopping the running internal-combustion engine can be performed much more quickly and with far lower vibrations with the belt-driven boost-recuperation machine. The fast response results in extended engine standstill times and thus further potential CO_2 savings.

The electrical energy now available to a greater extent can also be used efficiently for additional electrical loads/consumers. Essentially two paths can be followed. Firstly, it is possible to increase powerful loads/consumers from the 12 V to the 48 V level and thus also make them much more compact. Secondly, it is possible to use new 48 V loads/consumers whose application has to this point been prohibited to a large extent on the 12 V level by their high power demand. The advantages of the 48 V system with regard to the derived application cases are presented and discussed in detail in [3].

When the system properties of the boost-recuperation system are taken into account in the overall vehicle network, positive effects can be achieved in other areas. The operating strategy of a diesel engine in particular can according to [4] be influenced to the effect that the untreated NO_x emissions can be dramatically reduced by using the belt-driven system at non-stationary operating points. The achievable CO_2 benefit here is not significantly limited by the NO_x reduction.

Beyond reducing the untreated NO_x emissions, the application range of exhaust-gas treatment can be significantly broadened in the form of a NO_x storage catalyst (NSC) in the diesel drivetrain. Figure 7 shows the regeneration ranges of such a NO_x storage catalyst. The ranges of particularly high and particularly low loads prevent catalyst regeneration mode. The regeneration range can according to [4] be substantially broadened by imposing and relieving load on the internal-combustion engine by means of the boost-recuperation machine. This results in an increase in efficiency for the overall conversion of the storage catalyst and thus in cost-saving potential in the exhaust system of the diesel drivetrain. In SCR systems this can involve AdBlue consumption becoming apparent. In NSC systems the use of precious metals can be reduced and in special cases of combined NSC-SCR systems one of the systems can even be dispensed with.

Figure 7: Widening of the regeneration window for the NO_x storage catalyst
A Range of NSC regeneration.
1 Load decrease of boost-recuperation system (BRS),
2 Load increase with BRS.



References

[1] V. Denner: Zukunft gestalten – Innovationen für effiziente Mobilität. 34th International Vienna Motor Symposium, (2013).
[2] M. Uhl et. al.: Boost Recuperation Machine – Electric Motor for 48 V Systems, Technical Conference, Electric & Electronic Systems in Hybrid and Electric Vehicles and Electrical Energy Management, (2013).
[3] M. Uhl, M. Wüst, A. Christ: Electrified Powertrain at 48 V – More than CO_2 and Comfort. 22nd Aachen Colloquium Automobile and Engine Technology, 2013.
[4] M. Wüst et al.: Operating Strategy for Optimized CO_2 and NO_x Emissions of Diesel-Engine Mild-Hybrid Vehicles, 15th International Stuttgart Symposium, (2015).

Electric machines for motor-vehicle propulsion

The electric drive for vehicles with a hybrid or a purely electric drive is typically supplied by a high-voltage battery (see Traction batteries) with direct voltage, which an inverter converts into a symmetrical, three-phase alternating-voltage system (Figure 1). The voltage amplitude and the frequency of this alternating-current system can be changed. Figure 2 shows simplified characteristic torque, power, current and voltage curves obtained during inverter operation. In the base-speed range (between standstill and specific speed at which the maximum voltage is achieved in the alternating-voltage system) a constant torque is typically obtained. The range from the specific speed up to the maximum speed is called the field-weakening range. In this range the mechanical power of the drive is approximately constant, which corresponds to a torque drop proportional to the reciprocal of the speed. However, the behavior in the field-weakening range is actually dependent on the machine type (see Typical characteristic curves).

This review provides an overview of the most important features and influencing factors of electric machines for motor-vehicle propulsion (see also [1] and [2]).

Requirements

Figure 3 shows by way of example important operating points for purely electric vehicle operation and their importance to driveability. From these the contents of three of the four requirement groups listed in Figure 4 can be derived: Continuous operation, overload operation, and cycle efficiency. These groups and four others (strength, costs, installation space, and oscillation excitations) are explained in more detail in the following.

Figure 2: Simplified characteristic curves of an electric drive

1 Torque,
2 Mechanical power,
3 Line-to-line voltage U_{LL} (fundamental component),
4 Supply current I_{AC} (fundamental component).
A Base-speed range,
B Field-weakening range.

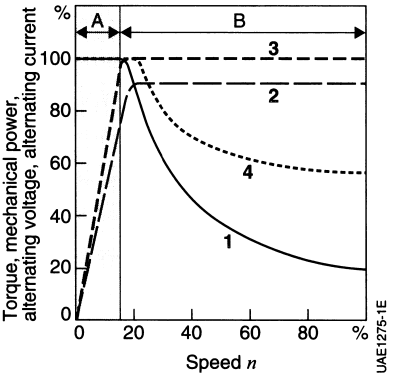
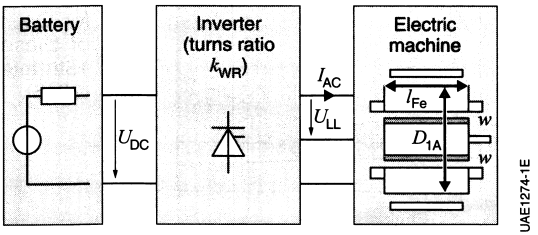


Figure 1: Electric drive

U_{DC} Battery voltage,
U_{LL} Line-to-line voltage,
I_{AC} Supply current from inverter to electric machine,
l_{Fe} Axial length of laminated core,
D_{1A} Outside diameter of stator,
w Number of turns.



Overload operation

Figure 5 shows typical characteristic torque and power curves of an electric vehicle drive. Overload operation characterizes driveability in accelerating and passing (overtaking) maneuvers. The electric machine cannot be continuously operated in overload operation as this would result in at least one limit temperature (for example, winding or magnet temperature) being exceeded. Typical requirement variables are the maximum torque at low speeds, the power at high speeds, and an associated minimum operating time. The characteristic curves of maximum torque and maximum power against speed are independent of the thermal properties of the machine and are called maximum limit characteristic curves. These are determined by the electromagnetic properties of the machine, the maximum supply current, and the battery voltage.

Continuous operation

Continuous operation characterizes driveability during longer trips at approximately constant driving speed and is a measure of the thermal properties of the machine. Continuous operation is defined by the reaching of a component limit tempera-

Figure 3: Important operating points for an electric motor-vehicle drive

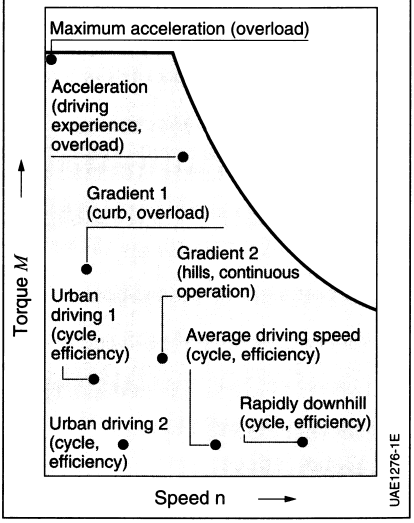


Figure 4: Requirement groups for an electric motor-vehicle drive

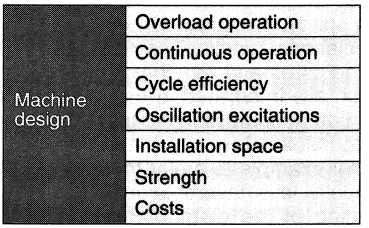
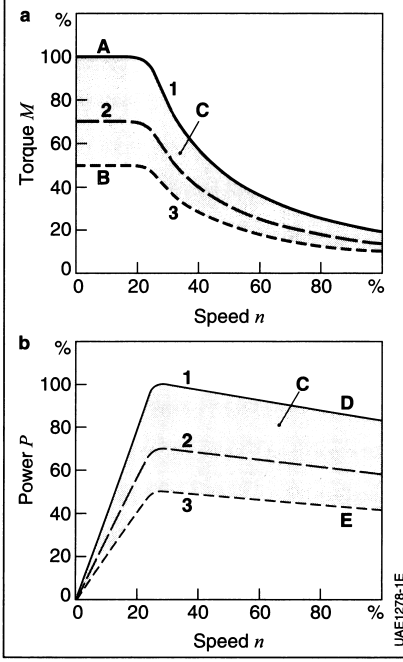


Figure 5: Typical characteristic torque and power curves for an electric motor-vehicle drive

a) Torque, b) Power.
1 Maximum limit characteristic curve,
2 60 s characteristic curve,
3 Continuous limit characteristic curve.
A Maximum torque, B Continuous torque,
C Overload operation,
D Maximum power at high speeds,
E Continuous power.

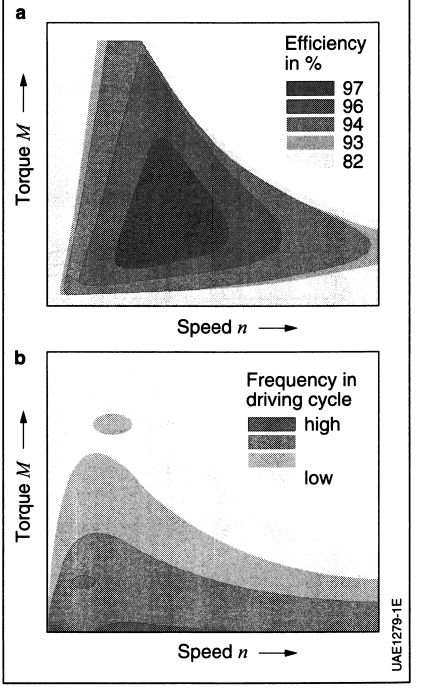


ture in thermal equilibrium. Typical requirement variables are the continuous torque at low speeds and a continuous power at high speeds. The characteristic curves of maximum continuous torque and maximum continuous power against speed are called continuous limit characteristic curves. Further characteristic curves with limited periods of time can be defined between the maximal limit characteristic curve and the continuous limit characteristic curve. Reaching the required period of time is dependent on the respective initial temperature.

Cycle efficiency

Figure 6 shows the efficiency map of an electric machine and the operating-point frequencies in a driving cycle. Depending on the distribution of the operating points in the driving cycle, other ranges are relevant to the cycle efficiency (typically rather small torques and medium speeds). Within the framework of design operating points below the maximum torque are as a rule calculated in such a way that optimum efficiency is obtained (regulation by the inverter, alternatively regulation to minimized oscillation excitations is also possible).

Figure 6: Efficiency map and operating-point frequencies for an electric motor-vehicle drive
a) Efficiency map,
b) Operating-point frequency in the driving cycle.



Oscillation excitations

In order to avoid unwanted sound radiation and service-life-reducing oscillations, oscillation excitations in the radial and tangential directions are calculated and minimized. The effects of torque ripple in the tangential direction depend on the mechanical systems connected to the machine (e.g. transmission). The requisite detailed knowledge of this system is often not available during the design process, which is why for the most part only a maximum peak-peak value of the torque-time characteristic is defined in the requirements.

In the specifications for electric machines as drives there are frequently maximum curves for the permissible sound-pressure or sound-power level. These can, depending on the quality of the mechanical model be calculated or estimated during the design process with the oscillation excitations in the radial direction.

Installation space

Especially for hybrid vehicles, which are still equipped with an internal-combustion engine as well as the electric drive, it is particularly important to minimize the machine size due to the small installation space available. The smaller this installation space, the greater the utilization of the electromagnetically active parts (laminated cores and winding). For machines which are integrated in the drivetrain of an internal-combustion engine (IMG, Integrated Motor-Generator), it may be sensible for example to prefer the electromagnetically less favorable individual-tooth winding of the distributed winding because the former saves a huge amount of

installation space thanks to its very short winding overhangs in the axial direction.

Strength
The electromagnetic-design requirements frequently conflict with the mechanical-design requirements. Narrow bars are often necessary in the core laminations to direct the torque-generating magnetic flux in the desired direction. These restriction can however become critical from a mechanical perspective if they are no longer able to withstand the loads that occur at high rotational speeds. A coupled optimization, which optimizes the combination of the two functions, is necessary here.

Costs
The development of electric machines as a motor-vehicle drive at Bosch is geared towards projects with high quantities (>100,000 units per year). In series production in this field even seemingly small cost savings per machine noticeably pay off overall. Because of the large number of competitors the target market price can only be achieved through massive cost optimization. The foundation for this is laid by a volume-production design with as little materials usage as possible. The rare-earth magnets in permanent-field synchronous machines are, on account of the high prices of the raw materials neodymium and dysprosium, a particularly important cost factor.

From the vehicle manufacturer's standpoint the system costs of the overall drive are to the fore (sum total of battery, inverter, and electric machine). In the system consideration it is sensible to optimize the components not individually and independently, but taking the overall system into account. This produces systems whose individual components are not necessarily optimum in terms of costs, but are so in terms of the overall system.

Types of electric machines

Table 1 provides an overview of the most important types of electric machines used in hybrid and electric vehicles. In hybrid vehicles a small installation space is for the most part decisive. This challenge can best be met with a permanent-field synchronous machine (PSM). The biggest drawbacks of this machine type are the relatively expensive rare-earth magnets in the rotor and the danger of these magnets being demagnetized.

Pure electric vehicles today also feature asynchronous machines (ASM, see Asynchronous machine) and electrically excited synchronous machines (ESM, see Synchronous machine). Asynchronous machines with squirrel-cage rotors are characterized by simple, robust construction, but often require rotor cooling due to the additionally occurring ohmic losses in the rotor. A very interesting alternative is the electrically excited synchronous machine, which on account of controllable rotor magnetization achieves particularly high efficiency levels and can manage on particularly little supply current. These advantages can lead to lower system costs and reduced energy consumption.

The switched reluctance machine (SRM, see Switched reluctance machine) is often examined as an alternative with a particularly simple rotor design, but has to date been unable to gain a foothold in any production vehicle due to the increased complexity of the power electronics (H-bridge instead of B6-bridge, see [1]) and its inherent susceptibility to noise generation.

Typical characteristic curves
These four types of electric machines essentially come into question for use as traction drives in electric and hybrid vehicles in view of general considerations. Ultimately, the classic machine types are favored: Asynchronous machine (ASM), permanent-field synchronous machine (PSM) and electrically excited synchronous machine (ESM) since these best fulfill the functional and economical requirements in the field of electrification. Which of these three machine types is the right solution in any concrete application

case can only be ascertained by means of a detailed comparison in which all three machines are designed according to the same requirements and the same installation space and electrical boundary conditions.

This comparison was conducted on the example of an electric vehicle (220 kW). The result is shown in the form of maximum limit characteristic curves with referred axis quantities in Figure 7. The abscissa represents the rotational speed referred to the maximum machine speed, the ordinate the mechanical drive power referred to the maximum available apparent electric power.

The differences between the three machine types on account of the principles involved can be clearly recognized after the maximum mechanical power is reached. The power of the ASM drops at a markedly greater rate in the direction of maximum speed. The reason for this is the leakage inductance in the stator and rotor, which significantly influences the operational performance in this range. Because of this performance the ASM is less suited to drive cases which also require high power at maximum speed. This is basically the case in expressway/interstate driving close to the vehicle's top speed.

The characteristic curve of the PSM is much flatter. Here the amount of magnet material used (for the most part neodymium iron boron, NdFeB, due to the high flux density of these magnets) and the manifestation of the reluctance torque of the magnetic circuit offers plenty of scope in the layout of the characteristic curve. The PSM is therefore highly suitable for universal use and is the preferred solution for drives with high power and torque densities, provided the expensive rare-earth magnets permit economic feasibility.

The ESM has two particularities distinguishing it from the ASM and the PSM: Firstly, the maximum achievable mechanical drive power is for the most part clearly above the other two machine types and, secondly, the characteristic curve runs almost horizontally up to the maximum speed. This is possible due to the additional controllability of the excitation current in the rotor. The ESM therefore best

Figure 7: Mechanical power for different types of electric machines
1 ESM: Electrically excited synchronous machine,
2 PSM: Permanent-field synchronous machine,
3 ASM: Asynchronous machine.

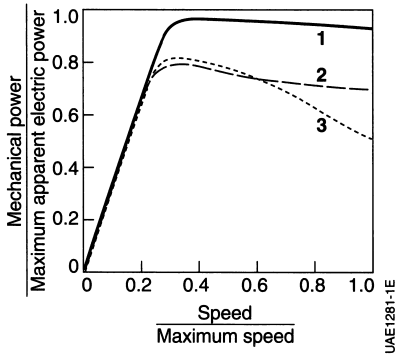


Table 1:Important machine types for electric motor-vehicle drives

Technology	PSM	ESM	SRM	ASM
Functioning principle	Synchronous	Synchronous	Synchronous	Asynchronous
Torque density	Very high	Very high	High	High
Improvement in efficiency at	High torque and low speed	High speed and low torque	High speed and low torque	High speed and low torque
Potential	Minimum installation space No rotor cooling	High efficiency Low supply-current demand	Very simple rotor design	Simple rotor design
Risks	Magnet price Demagnetization Short-circuit resistance	Current-transformer system, rotor Speed strength	Noise Outlay, power electronics	High current demand Rotor heating

converts the available electric power into mechanical power. Efficiency and power factor are accordingly higher for the ESM than for the ASM and the PSM. However, it is necessary to weigh up whether these advantages justify the design outlay for the excitation-current supply (collector rings or contactless transformers), and whether the problems of rotor cooling and of mechanical strength at higher speeds justify its use as a driving machine.

Aside from the immediate properties which are derived from the characteristic curve, further particularities are to be noted in the machine types: In both synchronous machines requirements with regard to limiting the steady-state and dynamic short-circuit current dramatically influence the design specifically at high speeds. Small short-circuit currents necessitate particularly with the PSM low power at maximum speed.

With the ASM much higher losses occur in the rotor. For this reason, active rotor cooling is for the most part necessary to achieve acceptable thermal conditions. Solutions with direct cooling of the short-circuiting rings in the rotor by oil or shaft cooling by water are possible here.

Essentially, the choice of the electric machine suitable for the concrete drive case is the result of an overall evaluation of technical and economic characteristics which cannot generally or readily be transferred from one application case to the other. This statement also reflects the fact that all three favored electric-machine types have been developed by different vehicle manufacturers to series maturity and are used in electrified drivetrains. A standard solution is not discernible at the present time.

Influencing factors on operational performance

The operational performance of the electric drive is influenced by the battery, the inverter, and the electric machine. Important influencing factors on the maximum limit characteristic curve of the asynchronous machine (ASM) and the permanent-field synchronous machine (PSM) are presented in the following. Diagrams are only shown for the PSM, since the ASM and the PSM perform similarly in most cases.

Figure 1 shows the electric drive consisting of battery, inverter and electric machine, and visualizes the influencing variables considered:
- battery voltage U_{DC},
- maximum supply current $I_{AC,max}$ from the inverter to the electric machine,
- axial length of the laminated core l_{Fe},
- number of turns w,
- outside diameter D_{1A} of the stator,
- turns ratio k_{WR} of the inverter.

Battery voltage

A higher battery voltage increases the maximum line-to-line voltage $U_{LL,max}$. This increases the specific speed of the drive (highest speed at maximum torque, Figure 8).

The battery voltage accordingly has no influence on the maximum torque in the base-speed range. The power in the field-weakening range on the other hand rises with the battery voltage. For the PSM the relationship is approximately linear and for the ASM approximately square.

For the maximum limit characteristic curves in Figure 8 it is presupposed that the battery can also provide the necessary power for the maximum limit characteristic curve. Otherwise the achievable power on the maximal limit characteristic curve is limited by the battery power.

Maximum supply current of the inverter

The maximum supply current has no influence on the power at maximum speed, since the current limit of the inverter is usually not reached here (Figure 9). The maximum torque in the base-speed range on the other hand rises with the maximum

supply current. Disregarding saturation, the relationship for the PSM is linear and for the ASM square. With a greater supply current the losses on the maximum limit characteristic curve also increase, above all in the base-speed range. This reduces the permissible overload time, i.e. the time until the electric machine's temperature limit is reached.

Figure 8: Influence of battery voltage U_{DC} on the maximum limit characteristic curve of a PSM

a) Torque curve and mechanical power,
b) Supply current and line-to-line voltage
1 Torque (a) and supply current (b) for U_{DC} = 230 V,
2 Torque (a) and supply current (b) for U_{DC} = 180 V,
3 Torque (a) and supply current (b) for U_{DC} = 130 V,
4 Mechanical power (a) and line-to-line voltage (b) for U_{DC} = 230 V,
5 Mechanical power (a) and line-to-line voltage (b) for U_{DC} = 180 V,
6 Mechanical power (a) and line-to-line voltage (b) for U_{DC} = 130 V.

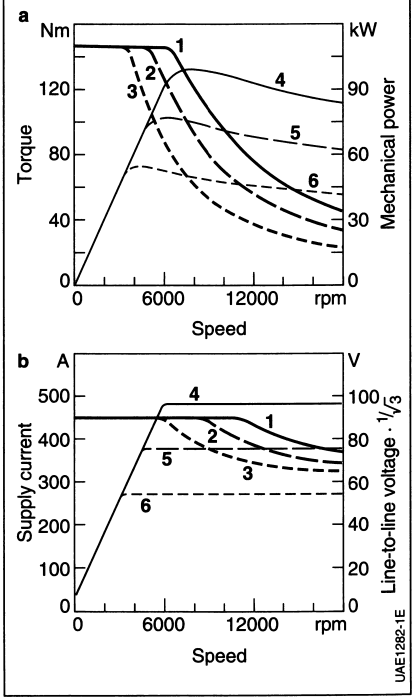
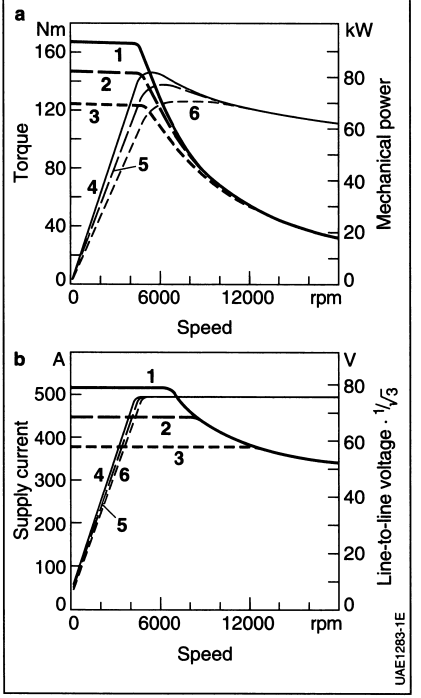


Figure 9: Influence of maximum supply current $I_{AC,max}$ on the maximum limit characteristic curve of a PSM

a) Torque curve and mechanical power,
b) Supply current and line-to-line voltage
1 Torque (a) and supply current (b) for $I_{AC,max}$ = 520 A,
2 Torque (a) and supply current (b) for $I_{AC,max}$ = 450 A,
3 Torque (a) and supply current (b) for $I_{AC,max}$ = 380 A,
4 Mechanical power (a) and line-to-line voltage (b) for $I_{AC,max}$ = 520 A,
5 Mechanical power (a) and line-to-line voltage (b) for $I_{AC,max}$ = 450 A,
6 Mechanical power (a) and line-to-line voltage (b) for $I_{AC,max}$ = 380 A.



Axial length of the electric machine's laminated core

The maximum torque in the base-speed range is proportional to the axial length of the laminated core (Figure 10). For the PSM the power remains virtually constant at high speeds. For the ASM on the other hand the power decreases at high speeds with increasing axial length. The cause is the increased leakage inductance.

Number of turns of the electric machine

If the number of turns is increased on the ASM or the PSM, the maximum torque in the base-speed range increase on ac-

Figure 10: Influence of axial length of the laminated core l_{Fe} on the maximum limit characteristic curve of a PSM

a) Torque curve and mechanical power,
b) Supply current and line-to-line voltage
1 Torque (a) and supply current (b) for 125 % l_{Fe},
2 Torque (a) and supply current (b) for 100 % l_{Fe},
3 Torque (a) and supply current (b) for 75 % l_{Fe},
4 Mechanical power (a) and line-to-line current (b) for 125 % l_{Fe},
5 Mechanical power (a) and line-to-line current (b) for 100 % l_{Fe},
6 Mechanical power (a) and line-to-line current (b) for 75 % l_{Fe}.

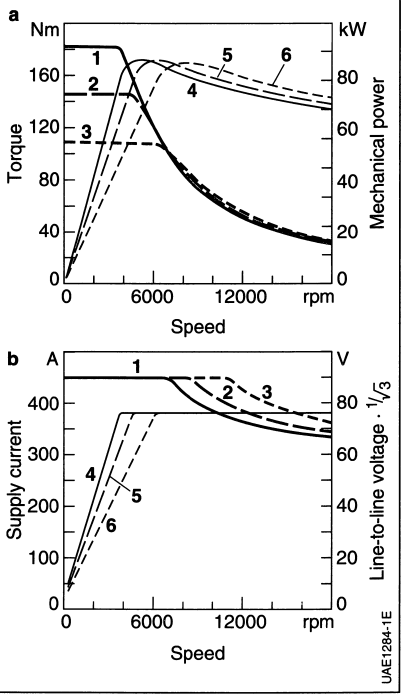
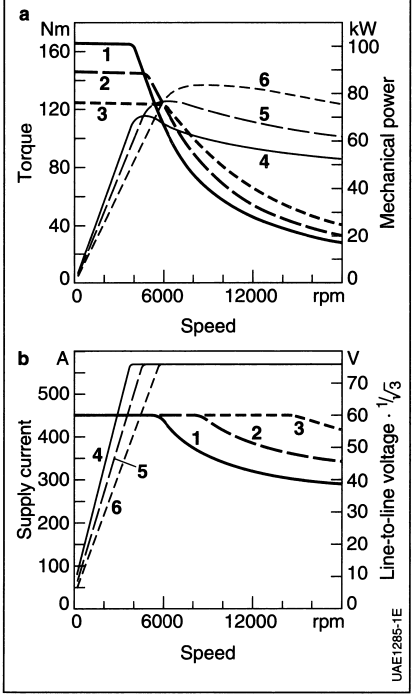


Figure 11: Influence of number of turns w on the maximum limit characteristic curve of a PSM

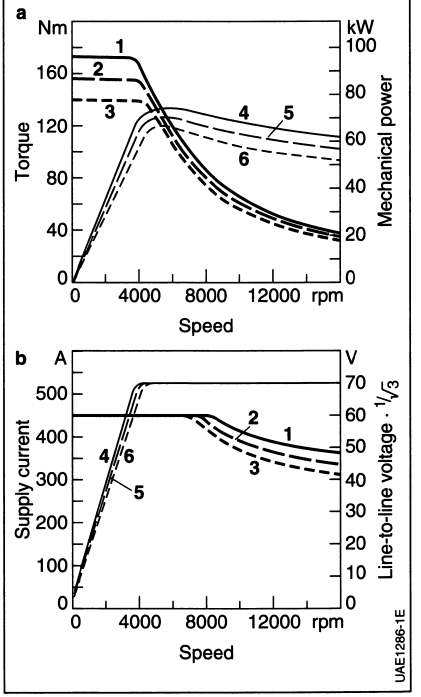
a) Torque curve and mechanical power,
b) Supply current and line-to-line voltage
1 Torque (a) and supply current (b) for 115 % w,
2 Torque (a) and supply current (b) for 100 % w,
3 Torque (a) and supply current (b) for 85 % w,
4 Mechanical power (a) and line-to-line current (b) for 115 % w,
5 Mechanical power (a) and line-to-line current (b) for 100 % w,
6 Mechanical power (a) and line-to-line current (b) for 85 % w.



count of the higher slot current linkage (Figure 11). The power at high speeds is on the other hand reduced, since the inductance of the electric machine is proportional to the square of the number of turns and thus the maximum line-to-line voltage is reached earlier. The slot's copper space factor (i.e. the ratio of copper space to slot space) is limited by the respective production engineering, therefore the conductor cross-section must be reduced if the number of turns is increased. In all, the winding resistance increases such that the losses on the maximum limit characteristic curve increase, above all in the base-speed range. This reduces the permissible overload time, i.e. the time until the electric machine's temperature limit is reached.

Figure 12: Influence of outside diameter D_{1A} on the maximum limit characteristic curve of a PSM

a) Torque curve and mechanical power,
b) Supply current and line-to-line voltage
1 Torque (a) and supply current (b) for D_{1A} = 200 mm,
2 Torque (a) and supply current (b) for D_{1A} = 180 mm,
3 Torque (a) and supply current (b) for D_{1A} = 162 mm,
4 Mechanical power (a) and line-to-line voltage (b) for D_{1A} = 200 mm,
5 Mechanical power (a) and line-to-line voltage (b) for D_{1A} = 180 mm,
6 Mechanical power (a) and line-to-line voltage (b) for D_{1A} = 162 mm.



Outside diameter of the stator

The influence of the stator outside diameter on the maximum limit characteristic curve is considered whereby the entire lamination section of the electric machine is linearly scaled with the outside diameter. In the base-speed range this produces a greater torque when the outside diameter is increased for the ASM and the PSM (Figure 12). For the chosen scaling it must be noted that the maximum supply current of the inverter and the number of turns of the electric machine have remained the same. With increasing outside diameter the current density decreases and the overload time increases.

Higher power is obtained for the PSM in the field-weakening range. For the ASM the power in the field-weakening range can also increase depending on design, but can also remain constant.

Turns ratio of the inverter

The turns ratio k_{WR} describes the relationship between the battery voltage U_{DC} and the maximum voltage $U_{LL,max}$ of the fundamental component (line-to-line voltage) of the alternating-current system:

$$U_{LL,max} = \sqrt{3}\, k_{WR}\, U_{DC} \quad \text{(eq. 1)}.$$

Different modulation techniques which give rise to different turns ratios can be used in regulation of the inverter. Common techniques and the associated ideal turns ratios are

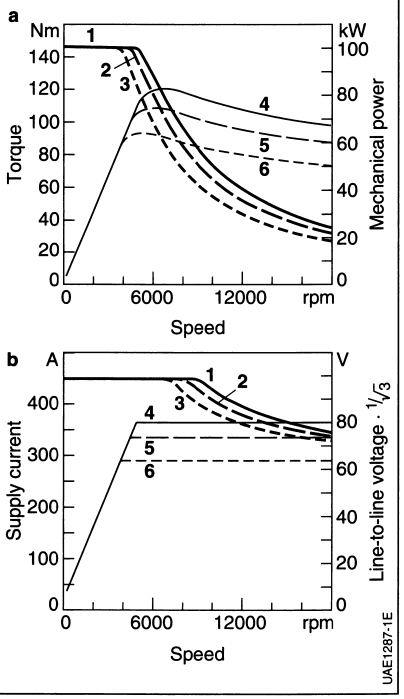
- sine-wave modulation, $k_{WR} = 0.3535$,
- space vector modulation, SVM; space vector pulse width modulation, SVPWM, $k_{WR} = 0.4082$,
- block clocking, $k_{WR} = 0.4502$.

Voltage drops via the power semiconductors are disregarded in the specified values.

Depending on the modulation technique used, the battery voltage is utilized differently and there are different voltage limits (Figure 13). k_{WR} therefore has no influence on the maximum torque in the base-speed range. With higher turns ratios the voltage limit is only reached at greater speeds so that higher power is achieved in the field-weakening range.

Figure 13: Influence of turns ratio k_{WR} on the maximum limit characteristic curve of a PSM

a) Torque curve and mechanical power,
b) Supply current and line-to-line voltage
1 Torque (a) and supply current (b) for $k_{WR} = 0.4502$,
2 Torque (a) and supply current (b) for $k_{WR} = 0.4082$,
3 Torque (a) and supply current (b) for $k_{WR} = 0.3535$,
4 Mechanical power (a) and line-to-line current (b) for $k_{WR} = 0.4502$,
5 Mechanical power (a) and line-to-line current (b) for $k_{WR} = 0.4082$,
6 Mechanical power (a) and line-to-line current (b) for $k_{WR} = 0.3535$.



Maximum apparent power

The maximum apparent power of the electric drive is

$$S_{max} = \sqrt{3}\, U_{LL,max}\, I_{AC,max} \quad \text{(eq. 2)}.$$

For a given modulation technique the following proportionality thus applies

$$S_{max} \sim U_{DC}\, I_{AC,max} \quad \text{(eq. 3)}.$$

The maximum apparent power is decisive in determining which maximum limit characteristic curve can be achieved with an electric drive.

The behavior of the electric drive when the battery voltage (influence on the power at maximum speed), the maximum supply current (influence on the maximum torque in the base-speed range) and the number of turns of the electric machine (exchange of torque in the base-speed range for power in the field-weakening range) are changed is described in the previous sections. In practice, however, these parameters are not adapted alone, but matched to each other by reference to the available maximum apparent power S_{max}. By varying the number of turns the same maximum limit characteristic curve can be set for any combinations of U_{DC} and $I_{AC,max}$ as long as S_{max} remains constant.

This is illustrated in a practical example in Figure 14. The starting position is an electric drive with the battery voltage U_{DC}, the maximum supply current $I_{AC,max}$,

and the number of turns w. This has the maximum limit characteristic curve 1 from Figure 14. If the battery voltage is now doubled, this produces the characteristic curve 2 from Figure 14 (similar to Figure 8); the maximum apparent power is doubled. To obtain the starting characteristic curve again, first the number of turns w must be doubled (characteristic curve 3 in Figure 14, similar to Figure 11) and then additionally the maximum supply current $I_{AC,max}$ halved (characteristic curve 4 in Figure 14, similar to Figure 9).

Adapting the number of turns produces in both cases (characteristic curve 1 and characteristic curve 4) the same maximum limit characteristic curve. Battery voltage and maximum supply current are indeed different, but the maximum apparent power is according to equation 3 constant. It must be noted that the number of turns can only be changed in discrete steps and therefore deviations can arise in reality.

Losses in electric machines

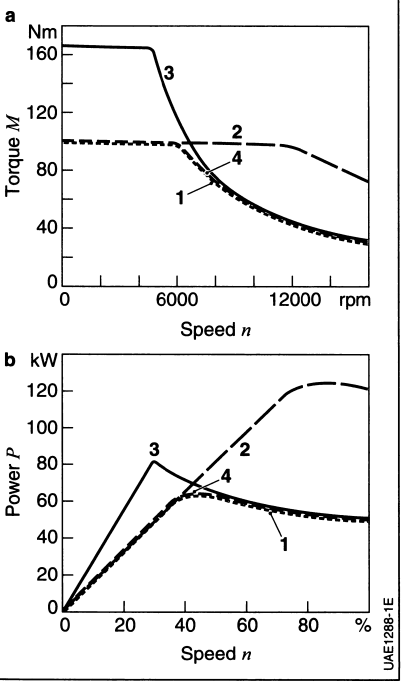
Figure 15 shows a schematic longitudinal section of an electric machine in which the most important assemblies are indicated and the associated types of loss are assigned to these.

Ohmic losses

Ohmic losses which in the first approximation are proportional to the square of the effective value of the current occur in the stator winding, in the rotor case of an asynchronous machine, and in the excitation winding of an electrically excited synchronous machine. For optimized de-

Figure 14: Setting the maximum limit characteristic curve of a PSM using the number of turns w at the same apparent power

1 Setting: U_B, $I_{AC,max}$, w.
2 Setting: $2\,U_B$, $I_{AC,max}$, w.
3 Setting: $2\,U_B$, $I_{AC,max}$, $2w$.
4 Setting: $2\,U_B$, $I_{AC,max}/2$, $2w$.



signs, current-displacement and proximity effects as well as parasitic circulating currents in parallel and delta connections must additionally be taken into account.

Iron losses

Iron losses occur due to rapid magnetic reversals predominantly in the laminated stator core of the machines. On account of the principles involved, synchronous and asynchronous machines exhibit absolutely no or only low-frequency sinusoidal components of flux through their basic fields in the rotor. The magnetic losses are usually broken down into two components:

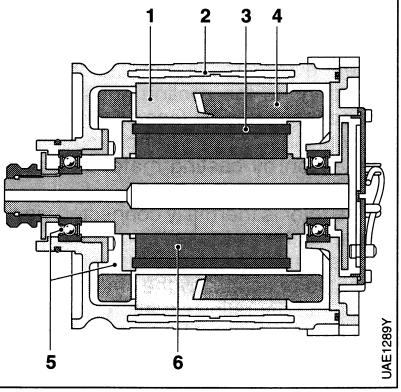
- hysteresis losses $\sim B^2 f$,
- eddy-current losses $\sim B^2 f^2$.

Here B denotes the flux density and f its frequency. Because of the high speeds and the accompanying high frequencies in automobile drives, the task of minimizing the eddy-current losses in particular poses a huge challenge. This is usually achieved by accordingly optimizing the lamination section and by using particularly thin core laminations (usually ≤ 0.35 mm) which are insulated against each other and stacked in the axial direction to form the laminated core.

Parasitic effects such as harmonic fields or harmonic components generated by the inverter can generate significant eddy-current losses both in the stator and in the rotor which must be taken into consideration in the design process.

Figure 15: Types of loss in an electric machine

1 Iron losses in the stator lamination,
2 Eddy-current losses in the housing,
3 Ohmic losses in the rotor cage (ASM), ohmic losses in the excitation winding (ESM), eddy-current losses in the magnets (PSM),
4 Ohmic losses in the stator winding,
5 Friction losses (bearing friction and windage),
6 Iron losses in the laminated rotor core.



Eddy-current losses

Further eddy-current losses can occur in the magnet blocks of a permanent-field synchronous machine and in structural parts (e.g. housings or end shields) due to the magnetic alternating fields generated by the machine. Of particular importance here are the eddy-current losses in the magnets, since these directly heat the magnetic blocks which are thermally for the most part poorly connected to the cooling system and consequently under high field load can lead to irreversible demagnetization.

Friction losses

The rotational movement of the rotor causes bearing-friction and windage losses, which on the one hand reduce efficiency and on the other hand can even result in unacceptable heating in the bearing itself. Particularly in hybrid vehicles, in which the electric machine always rotates but is not always actively in use, the friction losses have a significant influence on cycle efficiency.

Cooling of the electric machine

The losses are dissipated in the machine to cooling media. A powerful cooling system protects the individual components against overheating. Insufficient dissipation of the individual heat losses from the electric machine restricts the permissible continuous power. In overload operation the heat-storage capacity limits the time until a limit temperature is reached.

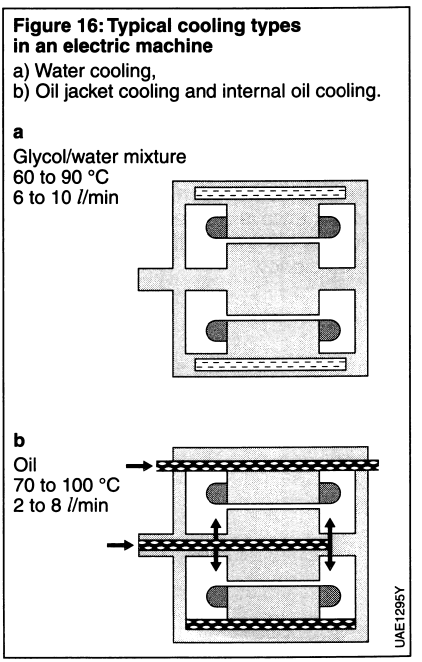
Cooling types

Different cooling media are provided on the vehicle side to cool the electric machine (Figure 16). The cooling effect is crucially determined not just by the cooling type, but also by the coolant temperature and the cooling mass flow.

Water cooling

The type most commonly encountered is water cooling, i.e. integration in a water/glycol circuit, which can also be used for the battery and the power electronics, and to heat the passenger compartment. Here the stator core is encased in a water-cooling jacket (Figure 16a).

Figure 16: Typical cooling types in an electric machine
a) Water cooling,
b) Oil jacket cooling and internal oil cooling.



Oil cooling

Oil cooling is advisable for integration in the manual transmission or final-drive differential. Cooling is also effected at the stator jacket (Figure 16b). But the rotor and the stator winding can also be cooled directly.

Air cooling

Air cooling is also used. Here air flows through a fan via large-area cooling fins along the stator jacket or even axially along the rotor.

Heat dissipation

Evaluation of the cooling of an electric machine requires knowledge of the individual, locally resolved losses and describes the heat paths to the cooling fluid. This is performed at the stator essentially by heat conduction from the winding to the cooling duct and at the rotor by convection. Mass node models and also CFD techniques (Computational Fluid Dynamics) are used for design purposes.

In the case of jacket cooling the main stator heat flow (90 %) proceeds from the winding via the stator-slot insulating paper to the stator core and from there to the housing with cooling duct. A specific role is assigned to winding cooling here. Electric insulation by paint, impregnating resin and insulating paper causes heavily directionalized thermal insulation. Cooling optimization here requires the interaction of electromagnetic design, production development, and thermal design.

Heat dissipation via the stator winding overhangs to the cooling housing by air can be increased by fans or additional oil cooling. Alternatively, a heat-conducting connection by casting materials can be established.

The rotor is thermally connected via the air gap to the stator and convectively connected via the side faces to the housing. It is heated or cooled by the stator, depending on the operating situation.

The distribution of losses in the electric machine changes with the speed (Figure 17). At low speed (top half of the pic-

ture) primarily ohmic losses are effective; at high speed (bottom half of the picture) iron losses are predominant. In this way, the different cooling paths work on the basis of speed.

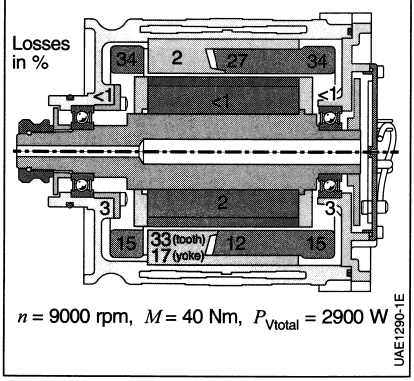
Thermal limitation

When the electric machine has a high power density, the maximum power is typically above the continuous power. Protection of the components against overheating is then ensured by a limiting function of the inverter's activation software. Limitation is performed typically for the limit temperatures of the stator and the rotor. In the case of the stator the insulation resistance limits the maximum winding temperature (e.g. 210 °C). Here the frequency distribution from vehicle operation is taken into account for the service-life evaluation. The rotor limit temperature is established for the permanent-field synchronous machine by the demagnetizing resistance of the magnets. This value must be derived from a worst-case analysis for the demagnetizing field at the limit temperature of the magnet (e.g. 160 °C).

Determination of the stator and magnet temperature during operation is subject to tolerances and requires a reduced control limit.

Figure 18 shows by way of example the continuous limit characteristic curve of a water-cooled permanent-field synchronous machine. Limitation for the stator temperature is effective here. Lowering the coolant temperature increases the drivable continuous torque.

Figure 17: Typical distribution of losses of an electric machine



References

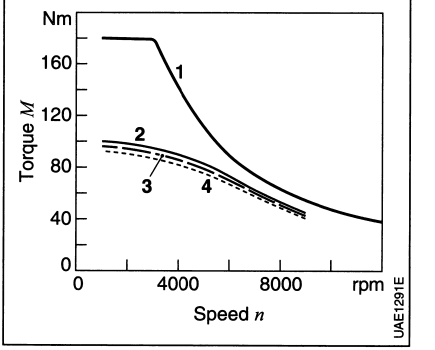
[1] K. Fuest, P. Döring: Elektrische Maschinen und Antriebe. 7th Ed., Vieweg+Teubner Verlag, 2007.

[2] A. Binder: Elektrische Maschinen und Antriebe. Springer-Verlag, 2012.

[3] R. Fischer: Elektrische Maschinen. 17th Ed., Carl Hanser Verlag, 2017.

Figure 18: Example of a continuous limit characteristic curve of an electric machine

Continuous limit characteristic curves with variation of the cooling-fluid temperature. Cooling-fluid throughflow is in all cases 6 l/min.
1 Maximum characteristic curve for torque,
2 Continuous torque for 65 °C,
3 Continuous torque for 75 °C,
4 Continuous torque for 85 °C.

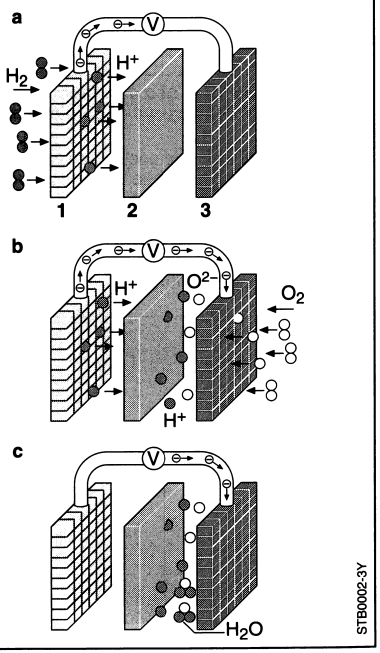


Fuel cells for the vehicle drive

Fuel cells are electrochemical converters which convert the chemical energy contained in the fuel directly into electrical energy. In a hydrogen/oxygen fuel cell, the hydrogen reacts with oxygen to create water, producing electrical energy in the process.

Fuel-cell drives are very appealing within the field of electromobility because they permit the accustomed convenience of conventional vehicles such as short refueling time and long range with expanded H_2 infrastructure. They allow for zero local emissions and cause little to no CO_2 emissions in the overall chain when using hydrogen from renewable energy sources.

Figure 1: Functioning principle of the PEM fuel cell
a) Hydrogen oxidation,
b) Oxygen reduction,
c) Water production.
1 Anode, 2 Membrane, 3 Cathode.



Functioning principle

A fuel cell consists of two electrodes (anode and cathode) which are separated from each other by an electrolyte. The electrolyte is not permeable for ions. The electrodes are connected to each other by way of an external electric circuit.

Polymer-electrolyte-membrane fuel cells (PEM-FC) are mostly used for mobile applications (Figure 1 and Figure 2). The functioning principle of the fuel cell is described in the following using this type.

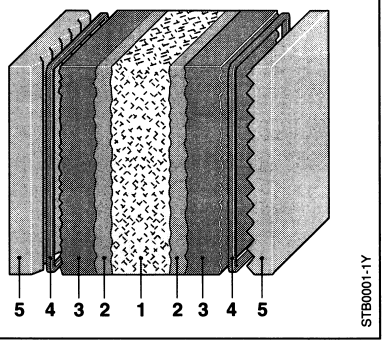
Functioning principle of the PEM fuel cell

In the PEM fuel cell, hydrogen is directed to the anode, where it is oxidized. H^+ ions and electrons are created (Figure 1a).

Anode: $2\,H_2 \rightarrow 4\,H^+ + 4\,e^-$.

The electrolyte is fashioned as a proton-conducting polymer membrane. It is permeable for protons, but not for electrons. The H^+ ions (protons), which are formed on the anode, pass through the membrane and reach the cathode. The polymer membrane must be sufficiently

Figure 2: Structure of the PEM fuel cell
1 Membrane (electrolyte layer),
2 Electrodes (catalyst layer),
3 Gas-diffusion layers,
4 Seals,
5 Bipolar plates.



moistened for this proton conduction. Air is directed onto the cathode – in automotive applications, to simplify matters, air is used with a oxygen content of around 21 %. Oxygen (O_2) is reduced at the cathode (Figure 1b). Reduction occurs using the electrons, which pass from the anode via the external circuit to the cathode.

Cathode: $O_2 + 4\ e^- \rightarrow 2\ O^{2-}$.

In a further reaction stage, the O^{2-} ions and the protons react to form water.

Cathode: $4\ H^+ + 2\ O^{2-} \rightarrow 2\ H_2O$.

The overall reaction of the fuel cell results in the transformation of hydrogen and oxygen into water (Figure 1c). Unlike an electrolytic-gas reaction, in which hydrogen and oxygen react with each other explosively, the transformation here takes place in a controlled reaction, since the reaction steps occur in separate spaces on the anode and cathode.

Overall reaction: $2\ H_2 + O_2 \rightarrow 2\ H_2O$.

The reactions described occur on catalysts in the fuel-cell electrodes. Platinum is frequently used for this.

The voltage of an individual hydrogen/oxygen fuel cell is theoretically 1.229 V at a temperature of 25 °C. This value is obtained from the standard electrode potentials (see Electrochemistry). This open-circuit voltage is however not reached during operation (Figure 3). Voltage losses can be put down, for example, to internal resistance or to gas-diffusion limitations. Essentially, the voltage is dependent on the temperature, the stoichiometric ratios of hydrogen and oxygen to the quantity of electricity produced, the hydrogen and oxygen partial pressures, and the current density.

In order to achieve the high voltages required for a technical application, individual cells are electrically connected in series to form a stack (Figure 4). These fuel cell stacks are made up of approximately 40 to 450 cells so that the maximum operating voltages are between 40 V and 450 V.

Stacks in the power range of 5 to 120 kW are typically used for automotive applications. High currents are achieved by appropriately sized membrane surfaces. Currents up to 500 A flow in automotive applications.

Figure 3: Electrical characteristic quantities of the fuel cell
1 Voltage,
2 Power density.

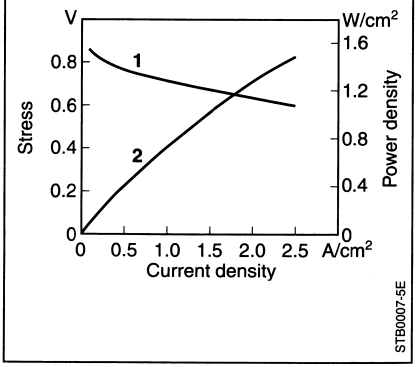
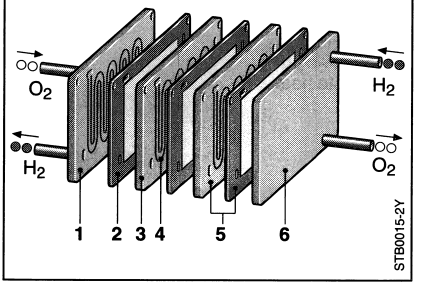


Figure 4: Structure of a fuel-cell stack
1 End plate,
2 Membrane electrode unit with membrane, catalyst material and seals,
3 Bipolar plates,
4 Gas channels,
5 Repeating units,
6 End plate.



Functioning principle of the fuel-cell system

Subsystems for oxygen and hydrogen supply and for thermal conditioning are required to operate a fuel-cell stack (Figure 5). There are, in principle, many ways of realizing these subsystems. The description given here corresponds to the state of the art of all major vehicle manufacturers.

Hydrogen supply

The hydrogen is stored in a 700 bar high-pressure tank. According to the state of the art, type-IV composite tanks with carbon-fiber casing and HDPE lining (high-density polyethylene) are used (diffusion barrier). The hydrogen is expanded to approximately 10 bar via a pressure reducer and fed to the anode via the hydrogen gas injector.

The metering valve (hydrogen gas injector) is an electrically actuated valve used to set the pressure of the hydrogen on the anode side. In contrast to fuel injectors in internal-combustion engines, the hydrogen gas injector must set continuous mass flows. Under certain conditions such as low H_2 demand or when a suction-jet pump is used the metering valve is also clocked or small pressure or mass-flow pulses are realized. A typical flow value for 100 kW is 2.1 g/s hydrogen. The pressure to be adjusted is max. 3 bar.

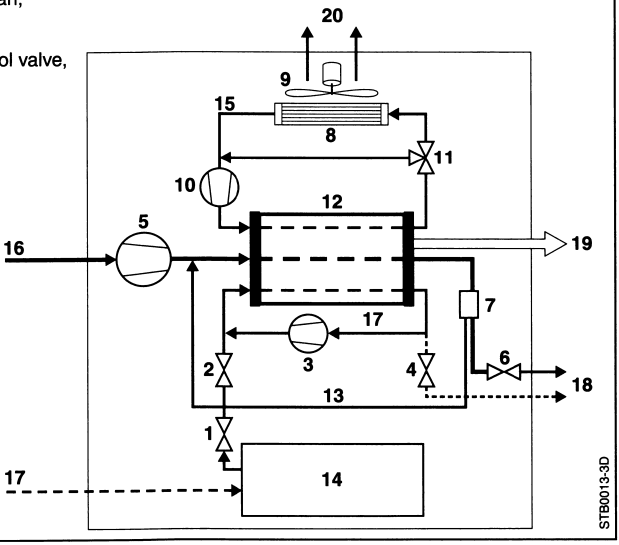
The anode side of the membrane has to be supplied with hydrogen everywhere equally. This is ensured by recirculating the hydrogen.

Disruptive foreign gases on the anode side are regularly removed via an electrically actuated valve, the drain valve. This prevents an enrichment of foreign gases from the fuel tank or of diffusion gases from the cathode side (e.g. nitrogen, water vapor). The valve is mounted on the anode side on the stack outlet. The valve open at zero current is also used to discharge unwanted quantities of water in the anode path.

The hydrogen necessarily discharged during draining is either heavily diluted with air or catalytically converted to water. Since these are only small amounts, the overall efficiency is only slightly reduced.

Figure 5: Fuel-cell system

1 Pressure reducer for hydrogen,
2 Hydrogen metering valve,
3 Hydrogen recirculation fan,
4 Drain valve,
5 Air compressor,
6 Dynamic-pressure control valve,
7 Water management,
8 Radiator,
9 Radiator fan,
10 Coolant pump,
11 Coolant control valve,
12 Fuel cell stack,
13 Water,
14 Hydrogen tank,
15 Coolant,
16 Air,
17 Hydrogen,
18 Exhaust gas,
19 Electric power,
20 Waste heat.



Oxygen supply
The oxygen required for the electrochemical reaction is taken from the ambient air. The required air mass of up to 100 g/s is inducted by a compressor in accordance with a desired fuel-cell power of 100 kW, compressed to max. 2.5 bar and fed to the cathode side of the fuel cell. The pressure in the fuel cell is adjusted by a dynamic-pressure control valve in the exhaust-gas path downstream of the fuel cell.

In order to ensure sufficient moisture for the polymer membrane, the supplied air is humidified using water management, either from the moist exhaust gases using an exchange membrane or by injecting condensed water.

Thermal management
Fuel cells have an electrical efficiency of approximately 60 %. This means that a large proportion of heat is released during the conversion of chemical energy. The heat must be dissipated. PEM fuel cells operate with an operating temperature of approximately 85 °C at a lower temperature level than internal-combustion engines. The radiator and radiator fan must therefore be designed to be bigger in fuel-cell systems in the vehicle drive in spite of the higher efficiency.

Because the coolant used is in direct contact with the fuel cells, it must be electrically non-conducting (deionized). The electric coolant pump directs a flow of coolant through all the incorporated components. The flow rate is up to 12,000 *l*/h. A coolant control valve splits the flow of coolant between the radiator and the radiator bypass.

Coolant based on a deionized water-glycol mixture is used. The coolant has to be deionized in the vehicle. The coolant flows through an ion exchanger filled with mixed-bed resin (not shown in Figure 5) and is cleaned by ion removal. The set-point value for coolant conductivity is less than 5 µS/cm.

Sensor technology
For fuel-cell systems specifically hydrogen-concentration sensors (see Hydrogen sensors) and coolant-conductivity sensors are required as new sensors. Furthermore, sensors in the hydrogen and air paths are subject to a heightened requirement with regard to hydrogen resistance and resistance to deionized water.

Efficiency of the fuel-cell system
In addition to quickly providing the requested power to the stack under the best possible operating conditions, it is important to design the system for primarily operating at high efficiency.

In Figure 6, the efficiency of a fuel-cell stack is compared with that of a fuel-cell system. The secondary loads (e.g. compressor) take up some of the electric power and thereby reduce the overall efficiency of the system. Fuel-cell systems demonstrate substantially higher efficiency than an internal-combustion engine, particularly in the frequently used part-load range.

Safety
Hydrogen is colorless and odorless. It can be ignited starting from a volume percentage of 4 % in air. Therefore, hydrogen-concentration sensors are placed at central positions in the vehicle (e.g. in the driver's cab, in the tank, in the engine compartment). These sensors are intended to help implement appropriate measures starting from a volume percentage of 1 %.

For the approval of hydrogen tanks the aim is to deliver maximum leakage rates of less than 0.1 g H_2 per day.

Functioning principle of the drivetrain

Expansion of the fuel-cell system with a battery

Fuel-cell vehicles are electric vehicles in which the energy for the electric drive is made available by the fuel-cell system. For several reasons, it is sensible to expand the drivetrain to include a traction battery:

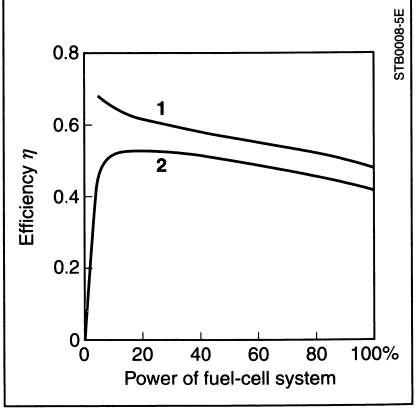
- The energy gained through recuperative braking can be temporarily stored
- The dynamic response of the drivetrain can be further increased
- The efficiency of the drivetrain can be further increased by shifting the load point between the fuel-cell system and the battery.

The topology with a fuel-cell system and traction battery corresponds to a series hybrid.

The ratio of battery power and total power varies for different applications. Typically, fuel-cell systems are used as the primary energy source in the drivetrain. Such vehicles are known as Fuel Cell Hybrid Vehicles (FCHV). Typically, fuel-cell systems have a power rating of 50 to 120 kW for the vehicle drive. Batteries have a power rating of up to 30 kW with an energy content of 1 to 2 kWh.

Figure 6: Efficiency of fuel-cell stack and fuel-cell system

1 Efficiency of fuel-cell stack,
2 Efficiency of fuel-cell system.



Alternatively, the battery can have a much greater power rating and energy content, and be recharged by the fuel-cell system where necessary. Fuel-cell power ratings of between 10 kW and 30 kW are sufficient for this. This drive concept is called Fuel Cell Range Extender (FC-REX).

Electrical-system typologies

One or more DC/DC converters spread the power between the fuel-cell system, the traction battery and the electric drive. In the main two different configurations are used, depending on whether the fuel-cell system or the battery is the primary source of drive energy. In systems with a large fuel-cell system and a small traction battery the DC/DC converter is situated ahead of the battery. In range-extender systems the power of the fuel-cell system is fed via a DC/DC converter to the battery circuit. In a third possible topology with two DC/DC converters, one for the fuel-cell stack and one for the traction battery, an optimized voltage is set independently of both. The latter offers advantages for the design of system components such as the compressor.

Components of the drivetrain

Electric drive

The electric drive consists of power electronics (a converter) and a synchronous or asynchronous machine that is energized by the converter to generate the required engine torque. Because the electric drive has a high power rating (approximately 150 kW), it is operated at voltages of up to 450 V; a vehicle manufacturer also selects a voltage up to 700 V. For safety reasons, the traction electrical system is isolated from the vehicle ground.

During braking the electric machine is switched to generator mode. In this process the vehicle's kinetic energy is converted into electrical energy and fed back to the traction battery.

The direct voltage of the traction electrical system is converted in the converter into a multi-phase alternating current where the amplitude is regulated as a function of the desired drive torque. Cyclically actuated IGBT output stages (Insulated Gate Bipolar Transistor) are the preferred choice for use as actuators.

Traction battery
Depending on the application, high-capacity or high-energy batteries are used with voltages of between 150 V and 400 V. Nickel-metal-hydride or lithium-ion batteries are used in high-capacity applications, while only lithium-ion batteries are used in high-energy applications. A battery-management system monitors the state of charge and the capacity of the battery.

DC/DC converter for traction battery
A DC/DC converter regulates the traction battery's charge and discharge currents where maximum currents of up to 300 A flow through the battery. This converter may be omitted in certain system configurations.

DC/DC converter for fuel cell
Another DC/DC converter regulates the current from the fuel-cell stack, where maximum currents of up to 500 A flow through the stack. This converter too may be omitted in certain system configurations.

DC/DC converter for 12 V electrical system
There is also – as in conventional vehicles – a 12 V electrical system for the low-power electric loads/consumers. The 12 V electrical system is supplied from the traction electric system. A DC/DC converter is used between the two systems for this purpose. This converter is electrically isolated for safety reasons. It operates unidirectionally or bidirectionally with a power output of up to 3 kW to keep the currents in the 12 V electrical system within a sensible range.

Outlook
Drives with fuel cells have already proven that they are suitable for everyday use.

Estimates from 2015 by the Department of Energy ([2], [3]) start out from fuel-cell-system costs of 54 US$ per kW at 80 kW and 500,000 vehicles per year and additionally 488 US$ per kilogram of hydrogen for the tank.

System simplifications deliver both cost and reliability advantages. One approach is to develop new polymer membranes for fuel cells where it is not necessary to humidify the reaction gases and which at the same time allows the operating temperature to be increased.

Vehicle models available on the market are the Hyundai ix35 Fuel Cell, the Toyota Mirai, and the Honda Clarity Fuel Cell (as at 2017). German automobile manufacturers are planning market launches in the coming years.

References
[1] D. Stolten, R.C. Samsun, N. Garland: Fuel Cells – Data, Facts, and Figures. 1st Edition, Verlag Wiley VCH, 2016.
[2] J. Marcinkoski, J. Spendelow, A. Wilson, D. Papageorgopoulos: DOE Hydrogen and Fuel Cells Program Record #15015 Fuel Cell System Cost.
https://www.hydrogen.energy.gov/pdfs/15015_fuel_cell_system_cost_2015.pdf
[3] B. James: Onboard Type IV Compressed Hydrogen Storage System Cost Analysis.
https://energy.gov/sites/prod/files/2016/03/f30/fcto_webinarslides_compressed_h2_storage_system_cost_022516.pdf

Emission-control legislation

Overview

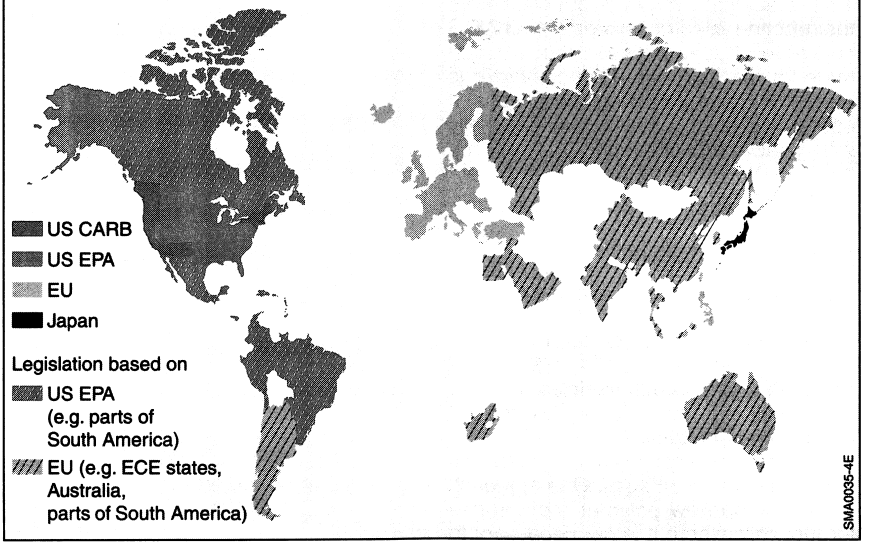
Provided for information purposes only, no liability assumed for completeness!

California introduced the first emission-control legislation for gasoline engines in the mid-1960s. These regulations became progressively more stringent in the ensuing years. In the meantime, all industrialized countries have introduced emission-control laws which define limits for gasoline and diesel engines, as well as the test procedures employed to confirm compliance. In many countries, regulations governing exhaust-gas emissions are supplemented by limits on evaporative losses from the fuel systems of vehicles with gasoline engines.

The most important legal restrictions on exhaust-gas emissions are listed below. Figure 1 provides an overview of the regions where the various types of legislation apply:

- CARB legislation (California Air Resources Board)
- EPA legislation (Environmental Protection Agency), USA
- EU legislation (European Union) and the corresponding UN/ECE regulations (United Nations/Economic Commission for Europe)
- Japanese legislation
- Chinese legislation

Figure 1: Overview of the regions where the various types of emission-control legislation apply
The map shows the regulations for passenger cars.



Classification
Countries with legal limits on motor-vehicle emissions divide vehicles into various classes:
- Passenger cars: Emission testing is conducted on a chassis dynamometer.
- Light commercial vehicles: Depending on national legislation, the top limit for gross vehicle weight is 3.5 to 6.35 t. Emission testing is performed on a vehicle chassis dynamometer (as for passenger cars)
- Heavy commercial vehicles: Gross vehicle weight over 3.5 to 6.35 t (depending on national legislation). Emission testing is performed on an engine test bench.
- Non-road (e.g. construction, agricultural, and forestry vehicles): Emission testing is performed on an engine test bench, as for heavy commercial vehicles.

Furthermore, there are emission regulations for other two-, three-, and four-wheel vehicles (see section Emission legislation for motorcycles), for locomotives, for boats and ships, and for non-road mobile machinery.

Test procedures
Japan and the European Union have followed the lead of the United States by defining test procedures for certifying compliance with emission limits. These procedures have been adopted in modified or unrevised form by other countries.

Legal requirements prescribe any of three different tests, depending on vehicle class and test objective:
- Type approval (TA) to obtain General Certification
- Random testing of vehicles from serial production conducted by the approval authorities (COP, Conformity of Production)
- In-field monitoring for checking the emission-reduction systems of privately owned production vehicles under real driving conditions

Type approval
Type approvals are a precondition for granting General Certification for an engine or vehicle type. For this purpose, test cycles must be driven under specific operating conditions and emission limits must be complied with. The test cycles and emission limits are specified individually by each nation.

Dynamic test cycles are specified for passenger cars and light commercial vehicles. The country-specific differences between the two procedures are rooted in their respective origins (see Test cycles for passenger cars and light-duty trucks):
- Test cycles designed to mirror conditions recorded in actual highway operation, e.g. the FTP test cycle (Federal Test Procedure) in the USA or the UN WLTC (World harmonized Light vehicle Test Cycle)
- Synthetically generated test cycles consisting of phases at constant cruising speed and acceleration rates, e.g. the MNEDC (Modified New European Driving Cycle) in Europe.

The mass of pollutant emissions from each vehicle is determined by operating it in conformity with speed cycles precisely defined for the test cycle. During the test cycle, the exhaust gases are collected for subsequent analysis to determine the pollutant mass emitted during the driving cycle (see Exhaust-gas measuring techniques).

For heavy-duty trucks and non-road applications, steady-state (e.g. 13-stage test) and dynamic test cycles (e.g. US HDDTC or ETC) are carried out on the engine test bench (see Test cycles for heavy commercial vehicles).

Serial testing (Conformity of Production)
Vehicle manufacturers normally perform serial testing themselves during production as part of quality control. The same test procedures and the same limits are generally applied as for type approval.

The approval agency audits the serial testing and can request retests. The strictest requirements are applied in the USA where very nearly 100 % quality monitoring is required (EPA: "Compliance Assurance Program" CAP; CARB: "Production Vehicle Evaluation" PVE).

In-field monitoring
In-field monitoring involves the detection of type-specific defects (e.g. design or manufacturing defects, defective maintenance instructions) which result in dramatically increased pollutant emissions during the operation of vehicles under normal conditions of use. To this end, a check is carried out in the field on production vehicles to ascertain that they comply 100 % with the emission regulations. Privately owned production vehicles are randomly chosen for testing. The mileage and age of the vehicle must be within defined limits. The emission-test procedure is partly simplified compared with type approval.

Emission-control legislation for passenger cars and light commercial vehicles

USA CARB legislation

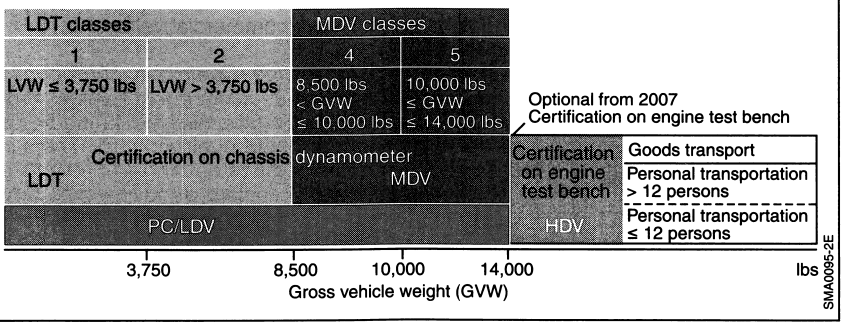
The emission limits of CARB (California Air Resources Board) for passenger cars and light commercial vehicles (light-duty trucks; LDT) are specified in the emission-control standards LEV I, LEV II, and LEV III (LEV, Low Emission Vehicle, i.e. vehicles with low exhaust and evaporative emissions). Since model year 2004 the LEV II standard has applied to all new vehicles up to a gross vehicle weight of 14,000 lbs (lb: pound; 1 lb = 0.454 kg, 14,000 lb = 6.35 t). Model year 2015 marked the beginning of the incremental introduction of the new emission level LEV III to be completed by 2025.

Originally the CARB legislation only applied in the US state of California, but has since been adopted in some other states.

Vehicle classes
Figure 2 shows an overview of the categorization of vehicles into vehicle classes.

Emission limits
CARB legislation specifies limits for the carbon monoxide (CO), nitrogen oxides (NO_x), non-methane organic gases (NMOG), formaldehyde (HCHO), and

Figure 2: Vehicle classes, CARB legislation
LDT Light Duty Truck, MDV Medium Duty Vehicle, HDV Heavy Duty Vehicle, PC Passenger Car, LDV Light Duty Vehicle, LVW Loaded Vehicle Weight (vehicle tare weight plus 300 lbs), GVW Gross Vehicle Weight.



particulate matter (diesel: LEV I and LEV II; gasoline: LEV III only) (Figure 3). On account of the extremely low emission limits, the LEV III legislation provides for an aggregate value of NMOG and NO_x.

Pollutant emissions are measured in the FTP 75 driving schedule (Federal Test Procedure). The emission limits are correlated with the route driven during the test and are expressed in grams per mile.

Within the period 2001 through 2004 the SFTP (Supplemental Federal Test Procedure) standard was introduced together with two other test cycles (SC03 and US06 cycles). There are also further limits that require compliance in addition to FTP emission limits.

Exhaust-gas categories

Automotive manufacturers are at liberty to deploy a variety of vehicle concepts within the permitted limits, providing they maintain a fleet average (see the section entitled "Fleet average"). The concepts are allocated to the following exhaust-gas categories, depending on their emission values for NMOG, CO, NO_x, and particulate emissions:

- LEV (Low-Emission Vehicle)
- ULEV (Ultra-Low-Emission Vehicle)
- SULEV (Super Ultra-Low-Emission Vehicle)

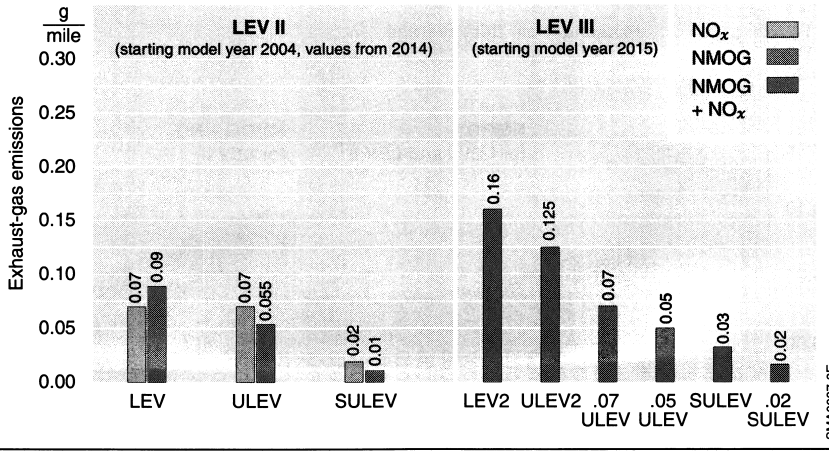
Since 2004 new vehicle registrations have been governed by the LEV II exhaust-emission standard. SULEV with significantly lower limits has been added. The LEV and ULEV categories will remain. The CO and NMOG limits from LEV I remain unchanged, but the NO_x limit is substantially lower for LEV II.

LEV III provides for a total of 6 vehicle categories (Figure 3), one of which is below SULEV. In addition to the categories of the LEV emission-control standards, three categories of zero-emission or partial zero-emission vehicles are defined (see ZEV program):

Phase-in

The LEV emission-control standards will not be introduced from one year to the next, but instead by means of a "phase-in", i.e. the incremental introduction of the requirements over a number of years for ever larger proportions of the new-car fleet, e.g. for LEV II 25 %/50 %/75 %/100 % of the newly registered vehicles in the model years 2004/5/6/7.

Figure 3: Exhaust-gas categories and emission limits for NO_x and NMOG of CARB legislation for passenger cars and light commercial vehicles



This is accompanied by a "phase-out" of the previous regulations.

The LEV III standard will be introduced between 2015 and 2025, phase-out of LEV II by 2019. Particulate emissions are covered by separate phase-in and phase-out from 2017 onwards.

Durability

In order to gain approval for vehicle types, vehicle manufacturers must certify that the emissions of limited pollutants do not exceed the respective limits over 50,000 miles or 5 years ("half useful life") and 100,000 miles (LEV I) or 120,000 miles (LEV II) or 10 years ("full useful life"). Manufacturers also have the option of certifying vehicles for 150,000 miles using the same limits that apply to 120,000 miles. The manufacturer then receives a bonus when the NMOG fleet average is defined (see the section entitled "Fleet average"). LEV III has basically extended durability to 150,000 miles.

Durability test

The vehicle manufacturer must supply two vehicle fleets from its production line for the durability test: One fleet in which each vehicle has driven 4,000 miles before the test, and one fleet for the endurance test with which the deterioration factors of individual pollutant components are measured.

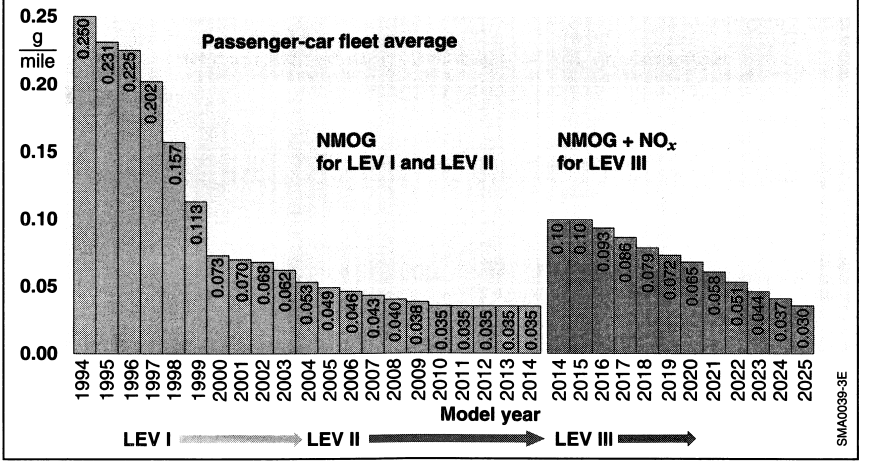
Endurance testing entails subjecting the vehicles to specific driving cycles over distances of 50,000 and 100,000 or 120,000 miles. Exhaust-gas emissions are tested at intervals of 5,000 miles and may not exceed the limits. Servicing and maintenance tasks may only be performed at the specified intervals.

Alternatively, the vehicle manufacturers may also use deterioration factors defined by legislation. To this end the vehicle manufacturer runs endurance tests and determines the percentage by which the emissions have deteriorated for each pollutant component. Then the factor by which the emissions from the new vehicle must be better than the legal limit so that the limits can still be maintained after the endurance run is defined.

Fleet average

Each vehicle manufacturer must ensure that its vehicles do not exceed an average specific emission limit for exhaust-gas emissions (Figure 4). The NMOG and NO_x emissions are used as the criteria in this regard with LEV III (with LEV I and LEV II only NMOG). The fleet average is calculated for LEV I and LEV II from the

Figure 4: Passenger-car fleet averages in comparison with NMOG standard



average of the NMOG limit for "half useful life" produced by all of the manufacturer's vehicles are sold within one year. The limits for "full useful life" apply to LEV III. The emission limits for the fleet average are different for passenger cars and light commercial vehicles.

The emission limit for the NMOG fleet average is reduced every year. To meet the lower fleet limit, manufacturers must produce progressively cleaner vehicles in the more stringent emissions categories in each consecutive year.

Fleet consumption (fuel consumption)
US federal lawmakers specify mandatory requirements on vehicle manufacturers with regard to the mean fuel consumption of their vehicle fleets, or the number of miles driven per gallon (federal law, applicable in all states, the relevant authority is the "National Highway Traffic Safety Administration", NHTSA). Here it is the metric "miles per gallon of fuel (mpg)" and not as otherwise usual in Europe for example the "amount of fuel used per unit distance" that is used. This representation of "fuel economy" corresponds to the reciprocal value of consumption per unit of distance.

The prescribed CAFE value (Corporate Average Fuel Economy) was until 2010 27.5 mpg for passenger cars. This equates to a fuel consumption of 8.55 *l* of gasoline per 100 km. Up to 2004 the value for light commercial vehicles was 20.7 mpg or 11.36 *l* of gasoline per 100 km. Between 2005 and 2010 fuel economy was raised each year by 23.5 mpg.

For model year 2011 the CAFE system was restructured for passenger cars and light-duty trucks (among others the definitions of passenger cars and light trucks) and ambitious target values were set: 33.3 mpg and 22.8 mpg for 2011, 40.1 mpg and 25.4 mpg for 2014, 43.4 mpg and 26.8 mpg for 2016, 46.8 mpg and 33.3 mpg for 2021, and 56.2 mpg and 40.3 mpg for 2025. For the stages 2012 through 2016 and 2017 through 2025 an EPA regulation harmonized with the CAFE regulation for limiting greenhouse emissions applies in parallel. The target value for a vehicle is dependent on the tire contact area between the wheels ("footprint").

At the end of each year the average "fuel economy" for each vehicle manufacturer is calculated based on the numbers of vehicles sold. The manufacturer must remit a penalty fee of $5.50 per vehicle for each 0.1 mpg its fleet falls short of the target. There is the possibility by way of a credit/debit system of balancing overfulfillment and underfulfillment in a year with credits and debits from other years.

Buyers will also have to pay a gas-guzzler tax on vehicles with especially high fuel consumption. Here, the limit is 22.5 mpg (10.45 *l* per 100 km). These penalties are intended to spur development of vehicles offering greater fuel economy.

The FTP 75 test cycle and the highway cycle are applied to measure CAFE fuel economy (see the section entitled "USA test cycles").

A Fuel Economy Label provides vehicle buyers with information on fuel consumption. Model year 2008 sees the introduction of "5 cycle fuel economy" (also called "5 cycle method"), which is intended to better reflect real driving conditions. This takes into account measurements in the SFTP schedules and in the FTP at –7 °C, which include among others aggressive acceleration, high final speed, and also operation with air conditioning.

Emission-free vehicles
The ZEV program ("Zero-Emission Vehicle") is to be used to force the development and market introduction of "zero-emission vehicles" ("technology forcing"), which are also intended to help reduce greenhouse emissions.

The ZEV program defines three categories of zero-emission and partial zero-emission vehicles. Genuine ZEV vehicles may not release any emissions when they are in operation. These vehicles are electric cars which are operated by a battery or a fuel cell.

PZEVs (Partial Zero-Emission Vehicles) are not emission-free, but they do emit particularly low levels of pollutants. They are weighted using a factor of more

than 0.2, depending on the emission-limit standard. The following requirements must be satisfied for the minimum factor of 0.2:
- SULEV certification for a durability of 150,000 miles or 15 years
- Warranty coverage extending over 150,000 miles or 15 years on all emission-related components
- No evaporative emissions from the fuel system (0 EVAP, zero evaporation). This is achieved by extensive encapsulation of the tank and fuel system. This results in greatly reduced evaporative emissions by the overall vehicle.

AT PZEVs (Advanced Technology PZEVs) are hybrid vehicles with a gasoline or diesel engine and electric motor and gas-powered vehicles (compressed natural gas, hydrogen).

The ZEV program envisages for the major automobile manufacturers minimum quantities for ZEV, AT PZEV and PZEV vehicles which will increase gradually between 2005 and 2017 and provide for ever greater proportions of AT PZEVs and ZEVs. The quantities are not calculated directly, but instead by means of so-called "ZEV credits", which depend among other things on the technology used, their performance, and the model year. From 2018 the PZEV category will be dropped and the necessary quantities of AT PZEVs and ZEVs will rise dramatically by 2025.

The relevant figures for the ZEV, AT-PZEV and PZEV emission-limit categories are 150,000 miles or 15 years ("full useful life").

In-field monitoring
Non-routine inspection
Random emission testing is conducted on in-use vehicles using the FTP 75 test cycle and – for gasoline vehicles – an evaporative-emission test. Depending on the relevant exhaust-gas category, vehicles with mileage readings below 90,000 or 112,500 miles are tested.

Vehicle monitoring by the manufacturer
Official reporting of claims or damage to specific emissions-related components and systems has been mandatory for vehicle manufacturers since model year 1990. The reporting obligation remains in force for a maximum period of 15 years, or 150,000 miles, depending on the warranty period applying to the component or assembly.

The reporting method is divided into three stages with an incremental amount of detail: Emissions Warranty Information Report (EWIR), Field Information Report (FIR), and Emission Information Report (EIR). The California Air Resources Board is notified of information regarding complaints, fault quotas, fault analysis, and impacts on emissions. The agency uses the Field Information Report as the basis to decide whether to enforce a recall action on the vehicle manufacturer.

USA EPA legislation

EPA (Environmental Protection Agency) legislation applies to all of the states where the more stringent CARB stipulations from California are not in force. The CARB regulations had already been adopted in 2004 in the north-eastern states of Vermont, Massachusetts, and New York; these were followed by further states.

The EPA legislation has been covered by the Tier 2 emission-control standard since 2004. The new Tier 3 emission-control standard will be introduced incrementally between 2017 and 2025. Tier 3 achieves an extensive harmonization of the requirements with the Californian LEV III program, e.g. the same certification categories, albeit with its own designation ("bin").

Vehicle classes

With the transition to Tier 2, an additional vehicle class has been introduced in the form of the MDPV (Medium-Duty Passenger Vehicle) (Figure 5). Thus, all vehicles up to a permissible total weight of 10,000 lbs (4.54 t) which are designed to carry up to 12 persons, are certified on a chassis dynamometer.

Light commercial vehicles are divided into two groups: LLDTs (Light Light-Duty Trucks) with a gross vehicle weight of up to 6,000 lbs (2.72 t) and heavier HLDTs (Heavy Light-Duty Trucks) with a gross vehicle weight of up to 8,500 lbs (3.86 t).

Since 2007 certification on a chassis dynamometer has also been optionally possible for vehicles weighing up to 14,000 lbs (6.35 t).

The classification of motor vehicles into the vehicle classes defined in Tier 2 has been retained for Tier 3.

Emission limits

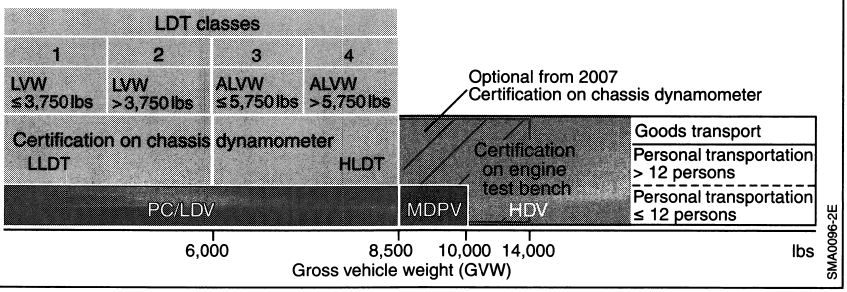
EPA legislation specifies limits for the pollutants carbon monoxide (CO), nitrogen oxides (NO_x), non-methane organic gases (NMOG), formaldehyde (HCHO), and particulate matter (PM). Pollutant emissions are measured in the FTP 75 driving schedule. The emission limits are correlated with the route driven during the test and are expressed in grams per mile.

The SFTP (Supplemental Federal Test Procedure) standard, comprising two further test cycles (SC03 and US06 cycles) has been in force since 2000. Prevailing emission limits must be met in addition to FTP emission limits.

Since the introduction of Tier 2 standards in 2004, vehicles with diesel and gasoline engines have been subject to identical exhaust-emission limits.

Figure 5: Vehicle classes, EPA legislation

LDT Light Duty Truck, MDV Medium Duty Vehicle, HDV Heavy Duty Vehicle, PC Passenger Car, LLDT Light Light-Duty Truck, HLDT Heavy Light-Duty Truck, MDPV Medium-Duty Passenger Vehicle, LDV Light Duty Vehicle, LVW Loaded Vehicle Weight, GVW Gross Vehicle Weight, ALV Adjusted Loaded Vehicle Weight (0.5 × tare weight + 0.5 × gross weight).



Exhaust-gas categories

Tier 2

For Tier 2, the limits were divided into 10 (for passenger cars and LLDT) and into 11 (for HLDT and MDPV) emission standards (bins) (Figure 6). For passenger cars and LLDT, Bin 9 and Bin 10 ceased to apply in 2007, and, for HLDT and MDPV, Bin 9 through Bin 11 ceased to apply in 2009.

The transition to Tier 2 has produced the following changes:

- Introduction of fleet averages for NO_x
- Formaldehyde (HCHO) is subject to a separate pollutant category
- Passenger cars and LLDT are treated identically to the greatest possible extent with regard to FTP limits
- "Full useful life" is increased, depending on the emission standard (Bin), to 120,000 or 150,000 miles

Tier 3

Seven selectable certification bins will furthermore be made available for Tier 3. However, the standards are now formulated both for the fleet values and for the bins as an aggregate value of NMOG and NO_x.

The Tier 3 legislation closely follows the Californian LEV III legislation in order to simplify the job of certification in the EPA and CARB areas for the vehicle manufacturers.

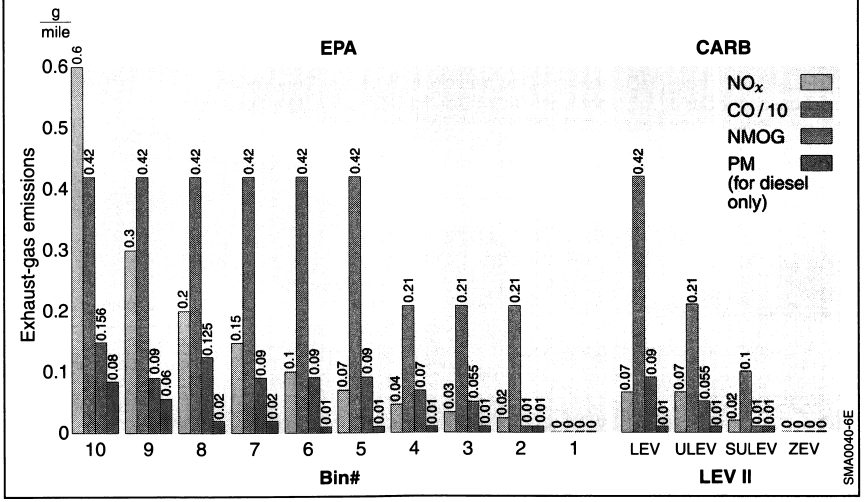
Durability has been extended for Tier 3 to 150,000 miles or 15 years compared with 100,000 miles or 10 years for Tier 2.

Tier 3 also covers heavy pick-ups and vans with a gross vehicle weight in excess of 6,500 lbs, which are now driven on a chassis dynamometer including the demanding SFTP test cycle. Within the Tier 2 legislation these vehicles were excluded from the SFTP tests.

Particulate-emission limits are prescribed with the Tier 3 legislation for each vehicle. A limit for the particulate number is currently not provided for.

The Tier 3 legislation is governed by the same fleet values for NMOG + NO_x as for LEV III.

Figure 6: Tier 2 emission limits for EPA legislation compared with CARB emission limits for LEV II.



Phase-in
At least 25% of all new passenger-car and LLDT registrations were required to conform to Tier 2 standards which took effect in 2004. The phase-in rule stipulated that an additional 25% of vehicles would be required to conform to the Tier 2 standards in each consecutive year. All vehicles have been required to conform to Tier 2 standards since 2007. The phase-in period for HLDT and MDPV terminated in 2009.

Tier 3 will be introduced between 2017 and 2025, phase-out of Tier 2 by 2019. Particulate emissions are covered by separate phase-in and phase-out from 2017 onwards.

Durability
The same criteria as for CARB apply to durability.

Fleet average
$NO2_x$ emissions are used to determine fleet averages for individual manufacturers under the EPA Tier2 legislation. Up to 2008 the value was 0.2 g/mile, since 2008 it has been 0.07 g/mile. With Tier 3 the NMOG+NO_x fleet average will be introduced as for CARB.

Fleet consumption (fuel consumption)
New vehicles registered in the area of responsibility are governed by the same CAFE regulations for determining fleet fuel consumption as for CARB (see CARB legislation, section "Fleet consumption").

For model years 2012 through 2016, in parallel with the CAFE legislation, EPA greenhouse-gas legislation was introduced, with target values of 34.1 mpg (CAFE) or 250 g CO_2 equivalent per mile (EPA) for 2016. In the period from 2017 through 2025 the target values will be further increased in stages, for EPA up to 54.5 mpg (163 g CO_2 per mile) and for CAFE up to 49.7 mpg (179 g CO_2 per mile).

In-field monitoring
Non-routine inspection
EPA legislation, just like CARB legislation, provides for an exhaust-gas emission test in accordance with the FTP 75 test method to be carried out on in-use vehicles on a random-test basis. Testing is conducted on low-mileage vehicles (10,000 miles, roughly one year old) and higher-mileage vehicles (50,000 miles, however at least one vehicle per test group with 90,000 or 105,000 miles, depending on the emission standard; roughly four years old). The number of vehicles is dependent on the number sold. For vehicles with gasoline engines, at least one vehicle per test group is also tested for evaporative emissions.

Vehicle monitoring by the manufacturer
For vehicles after model year 1972, the manufacturer is obliged to make an official report concerning damage to specific emission-related components or systems if at least 25 identical emission-related parts in a model year are defective. The reporting obligation ends five years after the end of the model year. The report comprises a description of damage to the defective component, presentation of the impacts on exhaust-gas emissions, and suitable corrective action by the manufacturer. The environmental authorities use this information as the basis for determining whether to issue recalls to the manufacturer.

EU legislation
The Directives of European emission-control legislation are proposed by the EU Commission and ratified by the Council of Environment Ministers and the EU Parliament. The basis of emission-control legislation for passenger cars and light commercial vehicles is Directive 70/220/EEC [1] from 1970. For the first time it defined exhaust-emission limits, and the provisions have been updated ever since.

The emission limits for passenger cars and light commercial vehicles LCV (light-duty trucks LDT) are contained in the emission-control standards Euro 1 (from 07/1992), Euro 2 (01/1996), Euro 3 (01/2000), Euro 4 (01/2005), Euro 5 (09/2009), and Euro 6 (09/2014).

Instead of a "phase-in" over several years as in the USA, a new emission-control standard is introduced in two stages. In the first stage newly certified

vehicle types must adhere to the newly defined emission limits. In the second stage – generally one year later – every newly registered vehicle (i.e. all types) must adhere to the new limits. Legislators can inspect production vehicles for compliance with emission limits (COP, Conformity of Production, and in-service conformity check).

EU directives allow tax incentives for vehicles that comply with upcoming exhaust-gas emission standards before they actually become law. Depending on a vehicle's emission standard, there are also a number of different motor-vehicle tax rates in Germany.

Vehicle classes

Until the Euro 4 legislation expired, vehicles with a permissible total weight below 3.5t were certified on chassis dynamometers; in this respect, a distinction was made between passenger cars (carrying up to nine persons) and light commercial vehicles (LCV) for goods transportation. There are three classes for LCV (Figure 7), depending on the vehicle reference weight (tare weight + 100kg). Engine certifications are performed for buses (carrying more than nine persons) and for vehicles with a permissible total weight in excess of 3.5t. LCV engines can optionally also be certified on an engine test bench.

When the Euro 5 and Euro 6 legislation comes into force, the vehicle reference weight (tare weight + 100kg) will be the distinguishing criterion with regard to the certification procedure. Vehicles with a reference weight of up to 2.61t will be certified on a chassis dynamometer. Vehicles with a reference weight in excess of 2.61t will be certified on an engine test bench. Flexible variations will be possible.

Emission limits

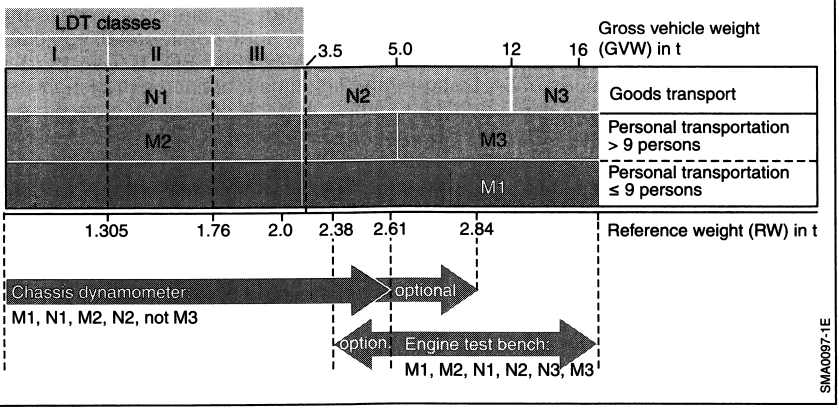
The EU standards specify limits for carbon monoxide (CO), hydrocarbons (THC, Total Hydro Carbons), nitrogen oxides (NO_x), and particulate matter (PM, for direct-injection gasoline engines as from Euro 5) (Figures 8 and 9).

The limits for hydrocarbons and nitrogen oxides for the Euro 1 and Euro 2 stages are combined into an aggregate value ($HC+NO_x$). Since Euro 3 a special NO_x limit in addition to the aggregate value has applied to diesel vehicles; in the case of gasoline vehicles, the aggregate has been replaced by separate HC and NO_x limits. The Euro 5 stage was

Figure 7: Vehicle classes, EU legislation

LDT Light Duty Truck, MDV Medium Duty Vehicle, HDV Heavy Duty Vehicle, PC Passenger Car, LDV Light Duty Vehicle, LVW Loaded Vehicle Weight, GVW Gross Vehicle Weight.

Note: The axes for GVW and RW must be considered separately!



introduced in two steps as Euro 5a and Euro 5b. Euro 5a (from September 2009) saw the addition for gasoline engines of an NMHC limit (non-methane hydrocarbons), Euro 5b (from September 2011) for diesel a particulate-number limit of $6 \cdot 10^{11}$ particulates per kilometer. This PN limit has also applied from Euro 6b (September 2014) and Euro 6c (September 2017) to vehicles with direct-injection gasoline engines (a higher interim value of $6 \cdot 10^{12}$ particulates/km for Euro 6b is possible on request).

The limits are defined based on mileage and indicated in grams per kilometer (g/km). Since Euro 3, emissions are measured on a chassis dynamometer using the MNEDC (Modified New European Driving Cycle).

From September 2017 the MNEDC will be replaced by the WLTC (Worldwide harmonized Light vehicles Test Cycle, 4 phases, see also Test cycles). The dynamic test cycle derived from road driving is supplemented by a strongly revised test procedure (WLTP) which is to make possible significantly more realistic con-

Figure 8: EU emission limits for gasoline-engine passenger cars and light commercial vehicles [3]

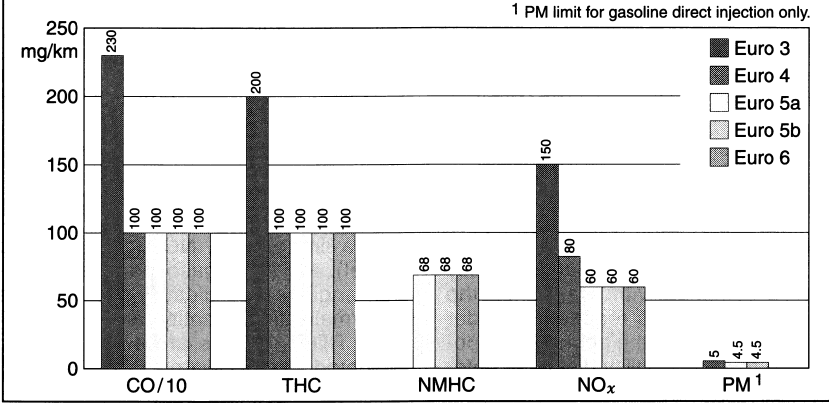
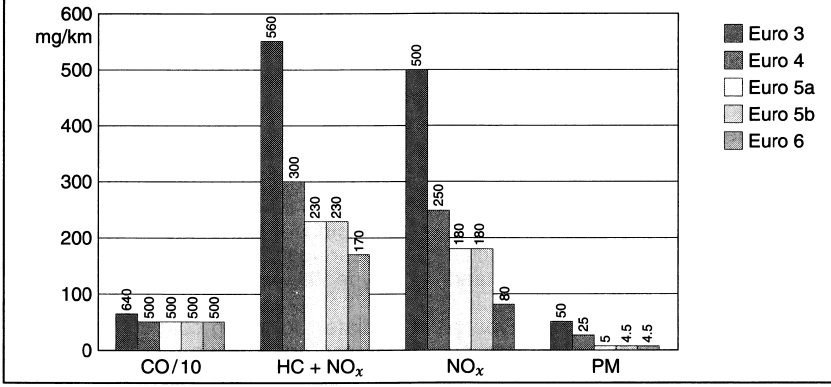


Figure 9: EU emission limits for diesel-engine passenger cars and light commercial vehicles [3]



sumption and emission measurements. Here the Euro 6c limits defined for the MNEDC are not adapted to the WLTC. This stage is called EU6d-temp since, aside from the change to the WLTC, the new "Real Driving Emissions" test (RDE) will also be introduced with a first stage. Euro 6d with the second RDE stage will apply from January 2020.

The limits are different for vehicles with diesel and gasoline engines, but with Euro 6 will be brought further into line.

The limits for LCV class 1 are the same as for cars. Passenger cars with a permissible total in excess of 2.5t were for Euro 3 and Euro 4 treated as LCVs and were therefore likewise categorized into one of the three LCV classes. This option ceases to apply from Euro 5.

Type approval

Type approval is conducted along similar lines to that in the USA, with the following differences: The pollutants HC, CO, NO_x and for diesel vehicles also the particulates and exhaust-gas opacity are measured. Test vehicles absolve an initial run-in period of 3,000 kilometers before testing. The limit values in the Type I test are "full useful life" limit values, i.e. they must be complied with even when the durability distance is reached. In order to take into account the aging of components up to the durability distance, deterioration factors are applied to the values measured in the type approval. These are defined in the legislation for every pollutant component; manufacturers are also allowed to present documentation confirming lower factors obtained during specified endurance testing (Type V test) over 80,000 km. From Euro 5 the endurance distance has been increased from 80,000 km to 160,000 km, where further alternative test procedures are possible.

Type tests

There are four essential type tests for type approval. Type I, Type IV, Type V and Type VI tests are used for vehicles with gasoline engines; only Type I and Type V tests are used for diesel vehicles. With Euro 6d-temp the RDE test Type Ia and the Ambient Temperature Correction Test are added for gasoline and diesel (ATCT, Type I test at 14 °C for converting the CO_2 emission from 23 °C to 14 °C, emission limits must be adhered to).

With the Type I test, the primary exhaust-gas test, the exhaust emissions are determined after cold starting in the MNEDC (Modified New European Driving Cycle) and since Euro 6d-temp in the WLTC (Worldwide Harmonized Light-Duty Vehicle Test Procedure). In addition, the opacity of the exhaust gas is recorded for diesel-engined vehicles.

The Type Ia test "Real Driving Emissions" (RDE) is intended to ensure that the emission limits are adhered to not only in the standardized cycle but also under real road conditions.

In the EU the RDE test is used both during the type approval and also in in-field monitoring. To this end a vehicle is equipped with a mobile measurement system (PEMS, Portable Emissions Measurement System, see Exhaust-gas measuring techniques) and driven for 90 to 120 minutes in normal road traffic. Various boundary conditions apply during this test (see Table 3, Test cycles). CO (not a limit value), NO_x (gasoline and diesel), PN (only gasoline direct-injection engines and diesel), and CO_2 (as a normalization quantity) are measured. Valid RDE journeys are evaluated and weighted with software tools (EMROAD and CLEAR). The average NO_x and PN emissions in the urban part and for the entire journey are compared with maximum values which are obtained from the product of emission limit multiplied by "Conformity Factor" CF. For NO_x CF will be introduced in two stages in 2017 2.1 and 2020 1.5; for PN 1.5 will apply from 2017. The CF values are to be adapted to the state-of-the-art.

Type IV testing measures evaporative emissions from parked vehicles. This primarily concerns fuel vapor that evaporates from the fuel-carrying system (fuel tank, pipes, etc.) (see also Exhaust-gas measuring techniques).

Type VI testing embraces hydrocarbon and carbon-monoxide emissions immediately following cold starts at –7 °C. Only the first section (urban section) of

the MNEDC is driven for this test. This test has been binding since 2002. A new test based on WLTC will be elaborated in 2017/18.

The Type V test assesses the long-term durability of the emissions-reducing equipment. In addition to the specified endurance test, alternative test procedures are possible from Euro 5 (e.g. test-bench aging).

CO_2 emissions

For CO_2 emissions there are no limits laid down by law until 2011; however, the vehicle manufacturers do enter into a voluntary commitment Europe. Because the manufacturers failed to reach their target, a fleet target value has been laid down by law for passenger cars. Within an introductory phase from 2012 through 2015 a fleet value of 130 g/km must be achieved (equating to 5.3 l/100 km gasoline or 4.9 l/100 km diesel). For light commercial vehicles a similar regulation stipulates a target value of 175 g/km for 2017. A further reduction for passenger cars to 95 g/km and for light commercial vehicles to 147 g/km has been decided on for 2020/2021. The target value for a vehicle is dependent on the vehicle tare weight.

From September 2017 new vehicle types must be certified in the WLTP (4 phases). The determined CO_2 values at 23 °C are corrected to 14 °C (by means of a factor determined in the Ambient Temperature Correction Test) and then converted with the aid of the CO2MPAS tool to MNEDC values, which will be used to check whether the targets are achieved up to and including 2020. On the basis of the relative achievement of the target values in 2020 and the underlying MNEDC and WLTP values a manufacturer-specific target value will be defined for 2021 and subsequent years.

Similarly to the US CAFE regulations, fines must be paid if the target value is not met, however there is no option of credit/debit balancing over a number of years.

In-field monitoring

EU legislation also calls for conformity-verification testing on in-use vehicles as part of the Type I test cycle. The minimum number of vehicles of a vehicle type under test is three, while the maximum number varies according to the test procedure.

Vehicles under test must meet the following criteria:

- Mileages vary from 15,000 km to 100,000 km, and vehicle age from 6 months to 5 years (from Euro 4).
- Regular service inspections were carried out as specified by the manufacturer.
- The vehicle must show no indications of non-standard use (e.g. tampering, major repairs, etc.).

If emissions from an individual vehicle fail substantially to comply with the standards, the source of the high emissions must be determined. If several vehicles display excessive emissions in random testing for the same reason, the results of the random test must be classified as negative. If there are various reasons, the test schedule may be extended, providing the maximum sample size is not reached.

If the type-approval authorities detect that a vehicle type fails to meet the criteria, the vehicle manufacturer must devise suitable action to eliminate the defect. The action catalog must be applicable to all vehicles with the same defect. If necessary, a recall action must be started.

Periodic emission testing

In Germany all passenger cars and light commercial vehicles and vans are required to undergo emissions inspections three years after their initial registration, and then at subsequent intervals of two years. For gasoline-engined vehicles, the main focus is on CO levels and λ closed-loop control, while for diesel vehicles, the opacity test is the main criterion. Data from the diagnostic system are taken into consideration for vehicles with On-Board Diagnostic systems (OBD).

Comparable tests are also available in other countries; in Europe, for example, in Austria, France, Spain, and Switzerland, and in many parts of the USA in the form of "Inspection and Maintenance".

Japanese legislation

The permitted emission values are also subject to gradual stages of severity in Japan. The limit values have since September 2007 been subject to further tightening within the framework of the "New Long Term Standards".

Diesel vehicles have since September 2010 been subject to further tightening of the limit values in the form of the "Post New Long Term Regulation". In the case of vehicles with gasoline engines, the previous synthetic test cycles are being replaced in two stages (2008 and 2011) by the more realistic JC08 cycle.

From 2018 Japan will also be introducing the WLTC based on UN GTR No. 15. Unlike the EU, only the first three phases apply and the NMHC limit values for gasoline-engine vehicles and the NO_x limit values for diesel are adapted.

Vehicle classes

Vehicles with a permissible total weight up to 3.5 t are essentially divided into three classes (Figure 10): passenger cars (up to ten seats), LDVs (Light-Duty Vehicles) up to 1.7 t, and MDVs (Medium-Duty Vehicles) up to 3.5 t. MDVs have higher limits for CO and NO_x emissions (for vehicles with gasoline engines) than the other two vehicle classes. For diesel engines, the vehicle categories are distinguished by NO_x and particulate limits.

Emission limits

The Japanese legislation specifies limits for carbon monoxide (CO), nitrogen oxides (NO_x), non-methane hydrocarbons (NMHC), particulate matter (for diesel vehicles, from 2009 also for gasoline direct injection with leaner NO_x 1 reducing technology) (Table 1 for gasoline engines and Table 2 for diesel engines).

Pollutant emissions are determined with a combination of 11-mode and 10•15-mode test cycles (see Japanese test cycle). Cold-start emissions are thus also taken into account. A new test cycle was introduced in 2008 (JC08). This replaces initially the 11-mode and as from 2011 also the 10•15-mode, so that only the JC08 will be used as the cold- and hot-start test. From 2018 the JC08 will be replaced by the WLTC.

Evaporative emissions

The exhaust-gas regulations in Japan include limits on evaporative emissions in vehicles with gasoline engines, which are measured using the SHED method (see Exhaust-gas measuring techniques). The WLTC-based test procedure of UN GTR 19 is expected to be introduced in 2020.

Durability

Manufacturers must demonstrate a durability of 45,000 km (New Long Term Standards) or 80,000 km (Post New Long Term Standards) for diesel vehicles.

Figure 10: Vehicle classes, Japanese legislation

MDV Medium Duty Vehicle, HDV Heavy Duty Vehicle, PC Passenger Car, LDV Light Duty Vehicle, GVW Gross Vehicle Weight, EIW Equivalent Inertia Weight.

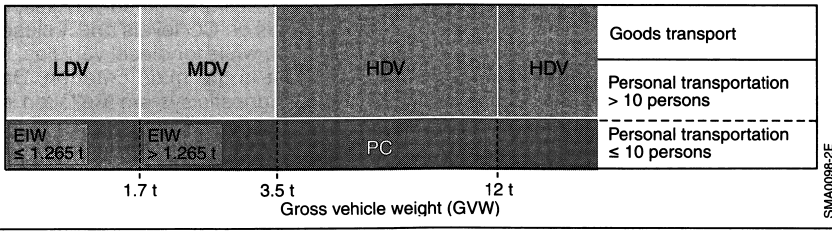


Table 1: Emission limits of Japanese legislation for gasoline-engine passenger cars

Vehicle category	Pass. cars, mini delivery vans, small commercial vehicles GVW ≤ 1.7 t		Medium-duty commercial vehicles 1.7 t < GVW 3.5 t	
Cycle Year (new types)	JC08 10/2009	WLTC 10/2018	JC08 10/2009	WLTC 10/2018
Limit values for gasoline engines				
NMHC [mg/km]	50	100	50	150
NO_x [mg/km]	50	50	70	70
CO [mg/km]	1,150	1,150	2,550	2,550
PM [mg/km] (direct-injection engines only)	5	5	7	7

Table 2: Emission limits of Japanese legislation for diesel-engine passenger cars

Vehicle category	Pass. cars, mini delivery vans, small commercial vehicles GVW ≤ 1.7 t		Medium-duty commercial vehicles 1.7 t < GVW 3.5 t	
Cycle Year (new types)	JC08 10/2009	WLTC 10/2018	JC08 10/2009	WLTC 10/2018
Limit values for diesel engines				
NMHC [mg/km]	24	24	24	24
NO_x [mg/km]	80	150	150	240
CO [mg/km]	630	630	630	630
PM [mg/km]	5	5	7	7

80,000 km apply to all stages for gasoline vehicles.

Fleet consumption
In Japan targets apply to the fleet consumption of a manufacturer for 2010 and 2015, based on target values for vehicle weight classes. For tax-incentive purposes (green tax program) there are two stages rewarding fuel consumption which is improved by 15% and 25% respectively.

Chinese legislation

China has, up to and including Stage 5 (applicable nationally from 2018), adopted the EU legislation to limit exhaust-gas emissions. Leading the way here was Beijing, introducing the stages early, followed by further metropolitan areas such as Shanghai and Guangzhou. With China 6a (7/2020) and 6b (7/2023) China is introducing its own legislation which for exhaust-gas emissions links the WLTC/P with Euro 6-based and US limit values and adds its own requirements.

For the Type I test the 4-phase WLTC including WLTP of GTR 15 is adopted. The limit values are fuel-neutral. For passenger cars and light commercial vehicles (class 1) for China 6a the following limits apply: for HC, NMHC, and NO_x the Euro 6 values for gasoline, for CO 0.7 · Euro 6 gasoline. PM (4.5 mg/km) and PN ($6 \cdot 10^{11}$ particulates/km) apply to diesel and gasoline (no exception for intake-manifold gasoline engines). A limit value for N_2O (laughing gas, a greenhouse gas) of 20 mg/km has been adopted from the USA. For China 6b the limit values for HC, NMHC, CO, and NO_x are reduced to approximately 50% of the Euro 6 gasoline values. PM is reduced to 3 mg/km, PN and N_2O remain the same. The durability requirement is 160,000 km for China 6a and 200,000 km for China 6b.

The Type VI test (at −7 °C) applies to gasoline and diesel as of China 6a. In the first two phases of the WLTC limit values

must be observed for HC of 1.2 g/km, CO of 10 g/m, and NO_x of 0.25 g/km.

China is also introducing an RDE test for the type approval and in-field monitoring. As of China 6a RDE tests must be conducted ("monitoring" of CO, NO_x, and PN, without limit values), as of China 6b "Conformity Factors" CF apply of 2.1 for NO_x and PN for all gasoline and diesel engines. The test procedures are based on an intermediate stage of the EU RDE legislation, e.g. the cold-start period is not taken into consideration and additional boundary conditions are defined, e.g. measurements up to an altitude of 2,400 m and lower top speeds for freeway driving. A review of the CF is envisaged by mid-2022.

The metropolitan areas can introduce China 6a and 6b early, but restrictions apply, e.g. before 7/2023 for RDE only monitoring and 160,000 km as the durability distance. For PN before 7/2020 $6 \cdot 10^{12}$ particulates/km apply.

China has also developed its own requirements for limiting fuel consumption and thus CO_2 emissions. In three stages (2005 through 2007, 2008 through 2015, and 2016 through 2020) limit values dependent on the vehicle weight in liters of fuel per 100 km apply to individual vehicle types. In parallel there is a Corporate Average Fuel Consumption (CAFC) regulation, which similarly to the EU CO_2 fleet targets sets vehicle-weight-dependent targets values for the consumption for the fleet of a manufacturer in the periods 2012 through 2015 and 2016 through 2020. Target values for the entire fleet are 6.9 *l* gasoline per 100 km for 2015 and 5.0 *l* gasoline per 100 km for 2020. For 2025 a target value of 4.0 *l* gasoline per 100 km is under discussion.

China uses a variety of measures, such as for example tax reductions, to promote the introduction of electrified vehicles (hybrids and plug-in hybrids, pure electric and fuel-cell vehicles). With the "New Energy Vehicle" legislation (NEV) manufacturers must from 2019 fulfill quotas for the sale of these vehicles (cf. CARB ZEV legislation).

Test cycles for passenger cars and light commercial vehicles

US test cycles

FTP 75 test cycle

The FTP 75 test cycle (Federal Test Procedure, Figure 11a) consists of speed cycles that were actually recorded in commuter traffic in Los Angeles. This test is also in force in some countries of South America and in Korea besides the USA (including California).

Conditioning

The vehicle is subjected to an ambient temperature of 20 to 30 °C for a period of 6 to 36 hours.

Collection of pollutants

The vehicle is started and driven on the specific speed cycle on a chassis dynamometer. The emitted pollutants are collected in separate bags during defined phases (see Exhaust-gas measuring techniques).

Phase ct (cold transient)

The exhaust gas is collected during the cold test phase (seconds 0 to 505 s).

Phase cs (cold stabilized)

The stabilized phase begins 506 seconds after start. The exhaust gas is collected without interrupting the driving cycle. Upon termination of phase cs, after a total of 1,372 seconds, the engine is switched off for a period of 600 seconds (hot soak).

Phase ht (hot transient)

The engine is restarted for the hot test. The speed cycle is identical to the cold transient phase (Phase ct).

Phase hs (hot stabilized)

For hybrid vehicles, a further phase hs is driven. It corresponds to the progression of phase cs. For other vehicles, it is assumed that the emission values are identical to the cs phase.

Assessment
The bag samples from the first two phases are analyzed during the pause before the hot test. This is because samples may not remain in the bags for longer than 20 minutes.

The sample exhaust gases contained in the third bag are also analyzed on completion of the driving cycle. The total result includes emissions from the three phases rated at different weightings.

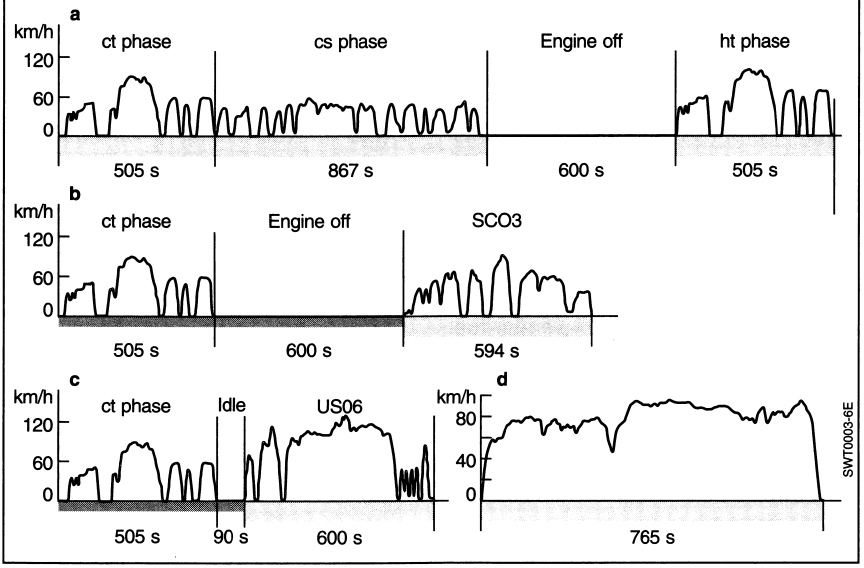
The pollutant masses of phases ct and cs are aggregated and assigned to the total distance of these two phases. The result is then weighted at a factor of 0.43.

The same process is applied to the aggregated pollutant masses from phases ht and cs, related to the total distance of these two phases, and weighted at a factor of 0.57. The test result for the individual pollutants (among others HC, CO, and NO_x) is obtained from the sum of the two previous results.
The emissions are specified as the pollutant emission per mile.

SFTP cycles
Tests according to the SFTP standard (Supplemental Federal Test Procedure) were phased in from 2001 to 2004. These are composed of two driving cycles, the SC03 cycle (Figure 11b) and the US06 cycle (Figure 11c). The extended tests are intended to examine the following additional driving conditions:

Figure 11: US test cycles for passenger cars and light commercial vehicles

	a	b	c	d
Test cycle	FTP 75	SC03	US06	Highway
Cycle distance:	17.87 km	5.76 km	12.87 km	16.44 km
Cycle duration:	1,877 s + 600 s pause	594 s	600 s	765 s
Mean cycle speed:	34.1 km/h	34.9 km/h	77.3 km/h	77.4 km/h
Maximum cycle speed:	91.2 km/h	88.2 km/h	129.2 km/h	96.4 km/h



- Aggressive driving
- Radical changes in vehicle speed
- Engine start and acceleration from a standing start
- Operation with frequent minor variations in speed
- Periods with vehicle parked
- Operation with air conditioner on

For preconditioning, the SC03 and US06 cycles proceed through the ct phase from FTP 75 without exhaust-gas collection. However, other conditioning procedures are also possible.

The SC03 cycle (for vehicles with air conditioning only) is carried out at a temperature of 35 °C and 40 % relative humidity. The individual driving schedules are weighted as follows:

- Vehicles with A/C systems:
  35 % FTP 75 + 37 % SC03
  + 28 % US06.
- Vehicles without A/C systems:
  72 % FTP 75 + 28 % US06.

The SFTP and FTP 75 test cycles must be successfully completed on an individual basis.

Cold-start enrichment, which is necessary when a vehicle with a gasoline engine is started at low temperatures, produces particularly high emissions. These cannot be measured in current emissions testing, which is conducted at ambient temperatures of 20 to 30 °C. An additional exhaust-gas test is performed at –7 °C on vehicles with gasoline engines in order to limit these pollutants. However, this test only prescribes a limit for carbon monoxide;a fleet limit was introduced for NMHC emissions in 2013.

Test cycles for determining fleet consumption

Each vehicle manufacturer is required to provide data on corporate average fuel economy. Manufacturers that fail to comply with the target values are required to pay penalties.

Fuel consumption is determined from the exhaust-gas emissions produced during two test cycles – the FTP 75 test cycle (weighted at 55 %) and the highway test cycle (weighted at 45 %). An unmeasured highway test cycle (Figure 11d) is conducted once after preconditioning (vehicle allowed to stand with engine off for 12 hours at 20 to 30 °C). The exhaust-gas emissions from a second test run are then collected. The CO_2 emissions are used to calculate fuel consumption.

Further test cycles

FTP 72 test cycle

The FTP 72 test routine – also known as the UDDS (Urban Dynamometer Driving Schedule) – corresponds to the FTP 75 test, but does not include the ht test component (hot test). This cycle is driven during the running-loss test for vehicles with a gasoline engine.

New York City Cycle (NYCC)

This cycle is also an element in the running-loss test (for vehicles with a gasoline engine). It simulates low speeds in urban traffic with frequent stops.

Hybrid cycle

For hybrid vehicles the phase hs (progression corresponding to phase cs) is attached to the FTP 75 cycle. This driving cycle thus corresponds twice to the UDDS cycle, which is why it is called 2UDDS.

European test cycle

MNEDC

The "Modified New European Driving Cycle" (MNEDC, Figure 12) has been in force since Euro 3. Contrary to the "New European Driving Cycle" (Euro 2), in which measurement of emissions only began 40 seconds after the vehicle was started, the MNEDC also includes a cold-start phase (including engine starting).

Conditioning

The vehicle is subjected to an ambient temperature of 20 to 30 °C for a minimum period of 6 hours. Since 2002 the starting temperature has been lowered to –7 °C for the Type VI test (only for vehicles with gasoline engines).

Collection of pollutants
The exhaust gas is collected in bags in two phases: the Urban Driving Cycle (UDC) at a top speed of 50 km/h and the Extra Urban Driving Cycle (EUDC, overland drive) at a top speed of 120 km/h.

Assessment
The pollutant mass measured by analyzing the bag contents is referred to the distance covered (see Exhaust-gas measuring techniques).

WLTC
Within the framework of the UN a new test cycle WLTC (Worldwide harmonized Light vehicles Test Cycle) was devised by the EU, Japan, India, and South Korea (among others) for passenger cars and light commercial vehicles and defined in UN GTR No. 15. The WLTC consists of four phases ("low", "mid", "high", and "extra high speed"). As an option, however, only the first three phases can be used. Further cycles were developed for special vehicle segments: for Japanese "kei cars" a toned-down variant of the WLTC and for the Indian market two cycles for vehicles with very low power-weight ratios ("low powered vehicle test cycles", LPTCs).

The dynamic test cycle derived from road driving is supplemented by a strongly revised test procedure (WLTP) compared with the previous regulation which is to make possible significantly more realistic consumption and emission measurements.

Among other things, the following have been defined for the WLTP test procedure: a method for shift-point calculation for manual transmissions, higher vehicle weights by taking into account payload and equipment options, improved methods for determining the coasting parameters, stepless adjustment of the chassis dynamometer, and 23 °C as the test temperature. Aside from the fundamental test procedure, special requirements for electrified vehicles such as hybrid and electric vehicles have also been elaborated.

Figure 13 shows the WLTC by way of example for class 3b, the class with the biggest-selling passenger cars and light commercial vehicles. Compared with the previous test cycle MNEDC (Figure 12),

Figure 12: MNEDC for passenger cars and light commercial vehicles
Cycle distance: 11 km.
Average speed: 33.6 km/h.
Top speed: 120 km/h.
UDC Urban Driving Cycle,
EUDC Extra Urban Driving Cycle.

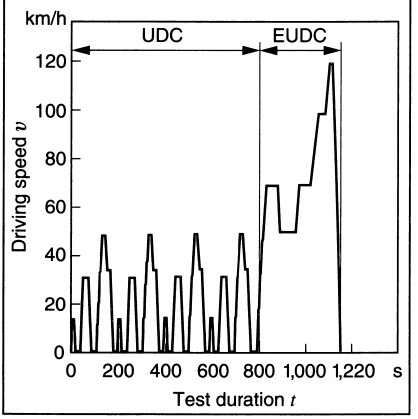
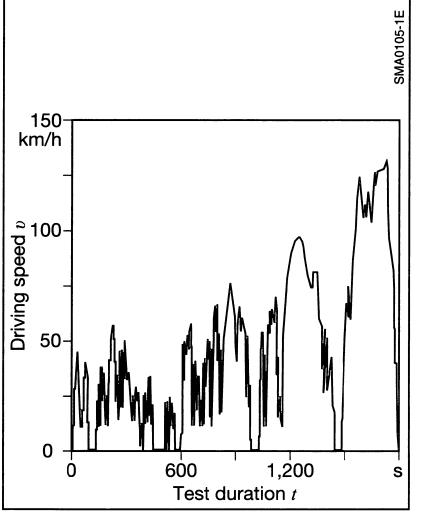


Figure 13: WLTC for passenger cars and light commercial vehicles



the changed dynamics of the WLTC can be clearly discerned.

RDE test

A new element of the legislation for passenger cars and light commercial vehicles is the RDE test ("real driving emissions"), which was developed in the EU and is being adopted by other countries like China and India (gasoline and diesel) and Korea and Japan (diesel only). The aim of the RDE test is to ensure that the emissions limits are observed not only in the standardized cycle but also under real road conditions. For the RDE test a vehicle is equipped with a mobile measurement system (PEMS, Portable Emissions Measurement System, see Exhaust-gas measuring techniques) and driven in normal road traffic. Various boundary conditions apply during this test, e.g. with regard to the proportions of city and intercity roads and freeway, the proportions of idle, of average speed, of minimum and maximum speeds, acceleration, cold start, and ambient conditions such as temperature and altitude. Table 3 sets out selected boundary conditions for the EU. RDE journeys can be made under conditions that are quite different from those in the Type I test (23 °C) thanks to the enhanced temperatures of −7 °C to +35 °C.

PEMS systems are available for CO, NO_x, PN, and CO_2. Valid RDE journeys are evaluated and if necessary weighted with software tools. The average emissions for a PEMS journey are compared with maximum values which are obtained from the product of emission limit multi-

Table 3: Boundary conditions for a valid RDE measurement run (selection, not all specified)

Parameter	Criterion		
Route distribution	Urban 34 % (29 to 44 %)	Interurban 33 % (±10 %)	Freeway 33 % (±10 %)
	In each case at least 16 km		
Speeds	$v \leq 60$ km/h v_D = 15 to 40 km/h (average speed)	60 km/h < $v \leq$ 90 km/h	90 km/h < $v \leq$ 145 km/h (up to 160 km/h for max. 3 % of the freeway time)
Total driving time	90 to 120 min		
Secondary loads	Operation of the air-conditioning system and the other secondary loads must correspond to their possible use by the load under normal driving conditions on the road		
Altimeter	Moderated: $h \leq 700$ m Extended: 700 m < $h \leq$ 1,300 m (measured values are divided by 1.6) Starting and end points may not differ in their height above sea level by more than 100 m		
Temperature	Moderated: Stage 1: 3 °C ≤ $T \leq$ 30 °C Stage 2: 0 °C ≤ $T \leq$ 30 °C Extended (measured values are divided by 1.6): Stage 1: −2 °C ≤ T < 3 °C, or 30 °C < $T \leq$ 35 °C Stage 2: −7 °C ≤ T < 0 °C, or 30 °C < $T \leq$ 35 °C		
Vehicle capacity weight and test weight	Basic capacity weight: driver, front passenger, test equipment Maximum weight from basic capacity weight and artificial capacity weight: max. 90 % of the sum total of the "weight of passengers" and the "capacity weight"		
Cold start	Cold-start period: until the coolant has reached 70 °C, max. 5 minutes cumulative combustion-engine operation Speed: average = 15 to 40 km/h, max. 60 km/h Emissions in the cold-start period are evaluated for the urban part and the total RDE journey with the normal methods		

plied by "Conformity Factor" CF. The CF takes into account the measuring inaccuracies which arise from measurements with PEMS conducted on the road compared with a measurement conducted on a chassis dynamometer.

Japanese test cycle

JC08 test cycle

In 2008 a new exhaust-gas test was introduced in the form of the JC08 (Figure 14), which initially replaced the 11-mode test as the cold test. Since 2011 only the JC08 has continued to be used (both as the cold-start and the hot-start tests). The cold test is weighted at 25 %, the hot test at 75 %. The pollutants are converted based on distance traveled, i.e. into grams per kilometer (g/km).

Emission-control legislation for heavy commercial vehicles

US legislation

Vehicle classes

Heavy commercial vehicles are defined in EPA legislation as vehicles with a gross vehicle weight over 8,500 lbs or 10,000 lbs (equivalent to 3.9 t or 4.6 t), depending on vehicle type (see Figure 5).

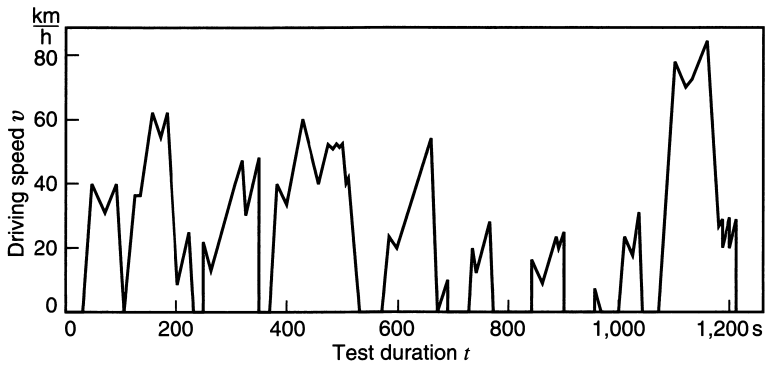
In California, all vehicles over 14,000 lbs (6.35 t) are classified as heavy commercial vehicles (see Figure 2). To a great extent, Californian legislation is identical to parts of EPA legislation. However, there is an additional program for city buses.

Emission limits

The US standards specify limits for diesel engines for hydrocarbons (HC), carbon monoxide (CO), nitrogen oxides (NO_x), particulate matter (PM), exhaust-gas opacity, and in parts for non-methane hydrocarbons (NMHC).

The permissible limits are related to engine power output and specified in g/hph (horsepower hour) (Figure 15 with values converted to g/kWh). The emissions are measured on the engine test bench during the dynamic test cycle

Figure 14: Japanese test cycle JC08 for passenger cars and light commercial vehicles
Cycle distance: 8.179 km.
Cycle time: 1,204 s.
Average speed: 24.5 km/h.
Top speed: 81.6 km/h.



with cold-starting sequence (HDDTC, Heavy-Duty Diesel Transient Cycle); the exhaust-gas opacity is measured using the Federal Smoke Test (FST).

New, more stringent regulations have applied to vehicles since model year 2004, with significantly reduced NO_x emission limits. Non-methane hydrocarbons and nitrogen oxides are grouped together in one aggregate (NMHC + NO_x). CO and particulate emission limits remain at the same level as model year 1998.

Another very drastic tightening of emission restrictions came into force in model year 2007. The NO_x and particulate emissions are separately limited and are a tenth of the previous values. This is not achievable without the use of emission-control systems (e.g. DeNOx catalytic converter, particulate filter, etc.).

A gradual phase-in will take place for NO_x and NMHC emission limits between model years 2007 and 2010.

To help compliance with severe particulate limits, the maximum permitted sulfur content in diesel fuel was reduced to 15 ppm from mid-2006.

For heavy commercial vehicles – in contrast with cars and light commercial vehicles – there are no limits specified for average fleet emissions and fleet consumption.

Consent decree

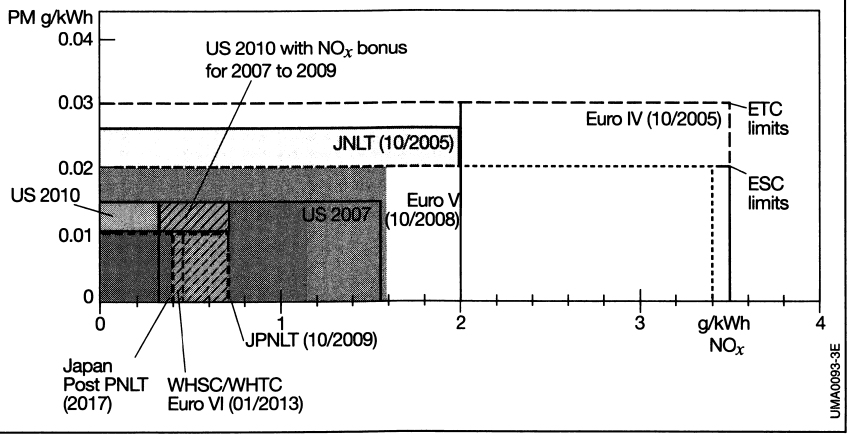
In 1998 a legal agreement was reached between EPA, CARB, and a number of engine manufacturers. It provides for sanctions against manufacturers if they make illegal modifications to engines to achieve optimized consumption in the highway cycle, resulting in higher NO_x emissions. The "Consent Decree" specifies that the applicable emission limits must also undercut the steady-state European 13-stage test in addition to the dynamic test cycle. Furthermore, emissions are not allowed to exceed the limits for model year 2004 by more than 25 %, regardless of driving mode within a specified engine-speed/torque range (Not-to-Exceed Zone).

These additional tests have been mandatory for all diesel commercial vehicles since model year 2007. However, emissions in the not-to-exceed zone may be up to 50 % above the emission limits.

Durability

Compliance with emission limits must be demonstrated over a defined mileage or a specific time period in which a distinction is drawn between three classes, each with increasing durability requirements:

Figure 15: Comparison of NO_x and PM emission limits for diesel fuel commercial vehicles: EU, USA, Japan [4]



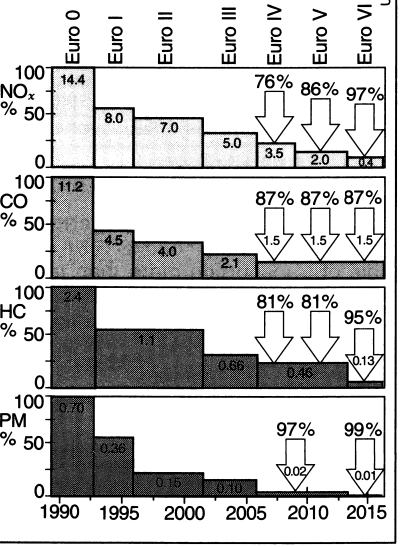
- Light commercial vehicles of 8,500 lbs (EPA) and 14,000 lbs (CARB) to 19,500 lbs: 10 years or 110,000 miles.
- Medium-duty commercial vehicles from 19,500 lbs to 33,000 lbs: 10 years or 185,000 miles.
- Heavy commercial vehicles over 33,000 lbs: 10 years or 290,000 miles.

Fuel-consumption requirement
In the USA there are separate requirements with regard to greenhouse-gas emissions (since 2014) and fuel consumption (since 2017).

EU legislation

Vehicle classes
In Europe, all vehicles with a permissible gross vehicle weight of over 3.5 t, or capable of transporting more than nine persons, are classified as heavy commercial vehicles (see Figure 7). The emission regulations (Euro standards) are based on the Basic Directive 88/77/EEC [2], which is continuously updated and extended.

Figure 16: EU emission limits for diesel commercial vehicles [4]
Values in g/kWh.
Percentage reductions refer to Euro 0.



Euro VI requirements are described by the created Directive 64/2012 EU [5]. Supplements to PEMS (Portable Emission Measurement System) and IUC (In Use Compliance) can be found for example in Regulation 2016/1718 EU [6].

Emission limits
As for passenger cars and light commercial vehicles, new emission-limit levels for heavy commercial vehicles are introduced in two stages. New engine designs must meet the new emission limits during type approval. One year later compliance with the new emission limits is a prerequisite for registering a new vehicle. The legislator can inspect conformity of production (COP) by taking engines out of serial production and testing them for compliance with the new emission limits.

For commercial-vehicle diesel engines, the Euro standards define emission limits for hydrocarbons (HC and NMHC), carbon monoxide (CO), nitrogen oxides (NO_x), particulates, and exhaust-gas opacity. The permissible limits are related to engine power output and specified in g/kW (Figure 16).

In October 2000 the limit stage Euro III was binding initially on new engine types and one year later on all newly manufactured engines. Emissions were measured during the 13-stage European Steady-State Cycle (ESC, see Test cycles for heavy commercial vehicles), and exhaust-gas opacity in the supplementary European Load Response (ELR) test. Diesel engines that are fitted with systems for exhaust-gas treatment (e.g. DeNOx catalytic converter or particulate filter) also had to be tested in the dynamic European Transient Cycle (ETC). These European test cycles were conducted with the engine running at normal operating temperature. Since Euro V a cold-start test has also been required.

Within the Euro III legislation a distinction was still made between large engines (displacement > 0.75 *l* per cylinder) and small engines (displacement < 0.75 *l* per cylinder and rated speed > 3,000 rpm), with higher permissible particulate-emission limits for the smaller units. This distinction is dropped for new certifications with the introduction of Euro IV.

There are separate emission limits for the ETC – for example, particulate limits are approximately 50 % higher than specified for the ESC because of the soot-emission peaks expected under dynamic operating conditions.

In October 2005 the Euro IV emission-limit level came into force initially for new type approvals, and for serial production one year later. All emission limits were significantly lower than specified by Euro III, but the biggest increase in severity applied to particulates, for which the limits were reduced by approximately 80 %. The following changes also applied after introduction of Euro IV:

- The dynamic exhaust-gas emission test (ETC) is obligatory – in addition to ESC and ELR – for all diesel engines.
- The continued functioning of emissions-related components must be documented for the entire service life of the vehicle (see Durability).

The Euro V emission-limit level was introduced in October 2008 for all new engine approvals, and one year later for all new serial-production vehicles. Only the NO_x emission limits were more severe compared to Euro IV.

In January 2013 the Euro VI emission-limit level for new engine types came into force (year later for all newly produced engines). Compared with Euro V, the nitrogen-oxide emissions are again reduced by 80 % and the particulate emissions by more than 60 % (referred to ETC limits for Euro V). New harmonized engine tests have been introduced with Euro VI. Here, too, there will be a stationary test (WHSC, World Harmonized Stationary Cycle) and a dynamic test (WHTC, World Harmonized Transient Cycle). Unlike the previous Euro V regulations, from Euro VI no specific particulate limits will be given for the transient test; the limits given will be identical to those for the stationary test.

Durability
Compliance with emission limits must be demonstrated over a defined mileage or a specific time period in which a distinction is drawn between three classes, each with increasing durability requirements:

- Light commercial vehicles up to 3.5 t gross vehicle weight (GVW): 6 years or 100,000 km (Euro IV and Euro V) and 160,000 km (Euro VI).
- Medium-duty commercial vehicles smaller than 16 t GVW: 6 years or 200,000 km (Euro IV and Euro V) and 300,000 km (Euro VI).
- Heavy commercial vehicles over 16 t GVW: 7 years or 500,000 km (Euro IV and Euro V) and 700,000 km (Euro VI).

Fuel consumption
Regulations for CO_2 emissions are currently in the planning stage; measurements are expected to be made with transportable measuring equipment on the vehicle while driving. CO_2 certification and type approval are expected in 2018 for semitrailer trucks and from 2020 for all heavy commercial vehicles.

Very low-emission vehicles
Prior to the introduction of the Euro V standard it was possible to meet voluntarily at an early stage more demanding, so-called EEV limit values (Enhanced Environmentally-Friendly Vehicle). Voluntary early meeting of these limits permitted tax incentives. EEV limit values for HC, NMHC, CO, and exhaust-gas opacity were lower than the Euro V limit values. NO_x and particulate limits were equivalent to the Euro V ESC limits. Natural-gas-powered city and tourist buses are typical examples of EEV vehicles.

Japanese legislation
Vehicle classes
In Japan, vehicles with a permissible gross vehicle weight of over 3.5 t, or capable of transporting more than ten persons are classified as heavy commercial vehicles (see Figure 10).

Emission limits
In October 2005 the "New Long Term Regulation" was introduced and was valid until the end of 2009. It stipulated emission limits for hydrocarbons (HC), nitrogen oxides (NO_x), carbon monoxide (CO), particulates, and exhaust-gas

opacity. Emission levels were measured in the newly introduced transient JE05 test cycle (hot test), and exhaust-gas opacity in the Japanese smoke test. The smoke test was conducted with three full-load operating points at 40.6 % and 100 % rated speed. With the introduction of the "Post New Long Term" emissions stage (09/09) there are now no smoke-test specifications.

The "Post New Long-Term Regulation" came into force in September 2009. The particulate and NO_x limits were reduced by almost two-thirds compared with 2005 levels.

Since October 2016 the "Post Post New Long Term" (Post PNLT, PPNLT) has applied to vehicles > 7.5 t with a renewed NO_x reduction to approximately 60 % of the previous PNLT requirements. The introduction of PPNLT will be completed in October 2018 also with limits for commercial vehicles ≤ 7.5 t. The deadlines at a glance:

- 10/2016 for heavy commercial vehicles > 7.5 t gross vehicle weight (not including semitrailer trucks),
- 10/2017 for heavy semitrailer trucks > 7.5 t gross vehicle weight,
- 10/2018 for vehicles with a gross vehicle weight of 3.5 to 7.5 t.

The harmonized test cycles WHSC and WHTC are used as the test cycle.

Durability
Compliance with the emission limits must be demonstrated over a defined mileage in which a distinction is drawn between three classes, each with increasing durability requirements:

- Commercial vehicles smaller than 8 t gross vehicle weight (GVW): 250,000 km.
- Medium-duty commercial vehicles smaller than 12 t GVW: 450,000 km.
- Heavy commercial vehicles over 12 t GVW: 650,000 km.

Fuel-consumption requirement
Fuel-consumption limits are prescribed for trucks and buses with a permissible gross vehicle weight of over 3.5 t. Two driving cycles (urban and extra-urban) are used.

Fuel consumption is determined on the engine test bench. Because consumption is very much dependent on the individual vehicle engine and equipment specification (e.g. drivetrain, rolling resistance, vehicle weight), the calculation is made using a conversion program. The requirements are for a commercial vehicle with a gross vehicle weight smaller than 10 t 7.4 km/*l*, for semitrailer trucks with a gross vehicle weight smaller than 20 t 2.9 km/*l*, and for buses with a gross vehicle weight smaller than 14 t 5.0 km/*l*. The figures apply to 2015.

Regional programs
In addition to the nationwide regulations for new vehicles, there are also regional requirements for the overall vehicle population aimed at reducing existing emission levels by replacing or upgrading old diesel vehicles.

The "Vehicle NO_x Law" has been in force since 2003 within, among other places, the greater urban area of Tokyo to vehicles with a gross vehicle weight of over 3,500 kg. It states that 8 to 12 years after a vehicle is first registered, the NO_x and particulate limits of the relevant preceding phase of emission limits must be adhered to (e.g. the 1998 limits from 2003). The same principle also applies to particulate emissions. Here, the regulation will already apply seven years after first vehicle registration.

Test cycles for heavy commercial vehicles

For heavy commercial vehicles, all test cycles are run on the engine test bench. In the transient test cycles, the emissions are collected and evaluated according to the CVS principle. The untreated emissions are measured in the stationary test cycles. Emissions are specified in g/kWh.

USA

Transient FTP dynamometer test cycle
Since 1987 engines for heavy commercial vehicles have been tested from a cold start on an engine test bench in a transient driving cycle (US HDDTC, Heavy Duty Diesel Transient Cycle). The test cycle is basically equivalent to operating an engine under realistic road-traffic conditions (Figure 17). It includes significantly more idle sections than the European ETC.

Federal Smoke Cycle
An additional test, the Federal Smoke Cycle, tests exhaust-gas opacity under dynamic and quasi steady-state conditions. In this test, for different load conditions (torque specifications), a number of sudden full-load accelerations are run on an engine test bench to determine exhaust-gas opacity.

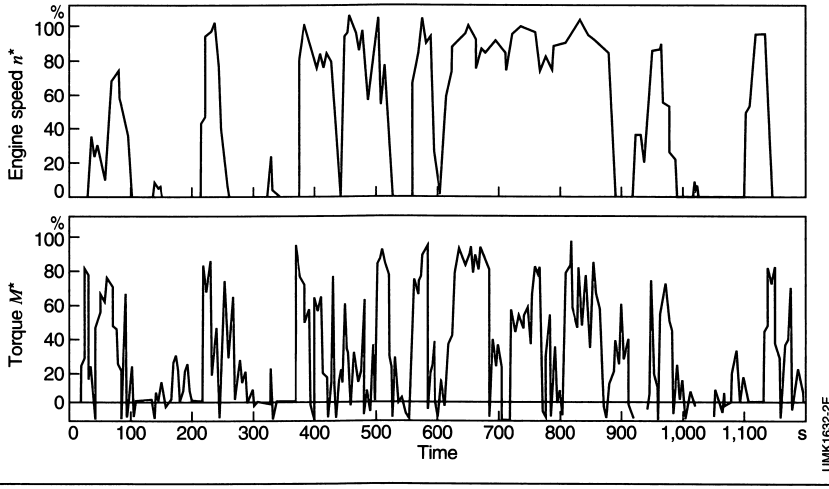
Additional test cycle
Since model year 2007 US emission limits have also had to comply with the European 13-stage test (ESC). Furthermore, emissions in the not-to-exceed zone (i.e. with any driving mode within a specified engine-speed/torque range) may be max. 50 % above the emission limits.

Europe

All European test cycles start with a hot engine.

European Steady-State Cycle
For vehicles over 3.5 t permitted gross vehicle weight and more than 9 seats, the 13-stage test ESC (European Steady-State Cycle, Figure 18), has been in force in Europe since the intro-

Figure 17: US FTP (Heavy-Duty Diesel Transient Cycle, HDDTC) for heavy commercial-vehicle engines
Both the nominal engine speed n^* and the nominal torque M^* are taken from tables specified by legislation.



duction of Euro III (October 2000). The test procedure specifies measurements in 13 steady-state operating states calculated from the engine full-load curve. The emissions measured at each operating point are weighted according to certain factors. This also applies to power output. The test results are obtained for each pollutant by calculating the total of the weighted emissions divided by the total of the weighted power output.

An additional three NO_x tests may be performed in the test range when certification is performed. The NO_x emissions may not vary by a significant degree from the levels measured at the adjacent operating points. The additional measuring has the goal of preventing engine modifications performed specially for the test.

Figure 18: Steady-state 13-stage test ESC (Europe)

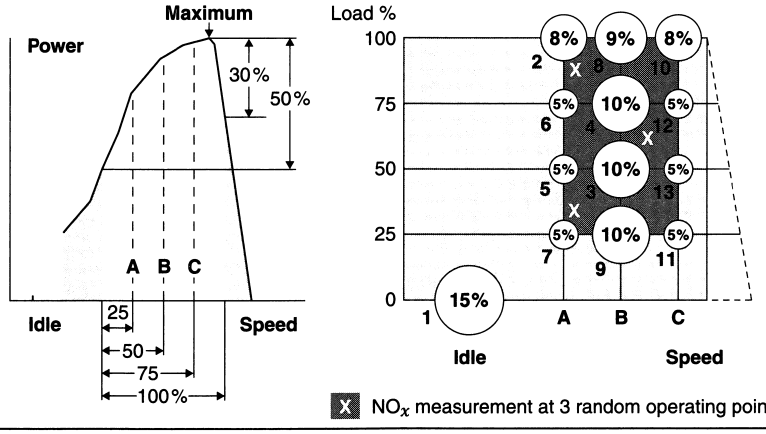
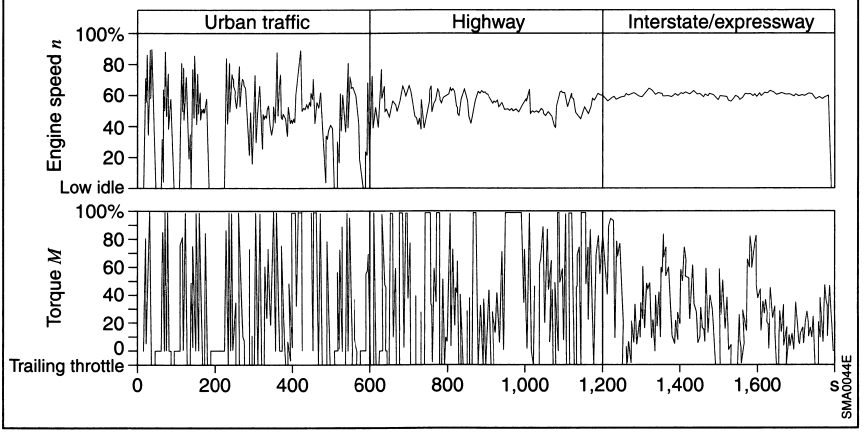
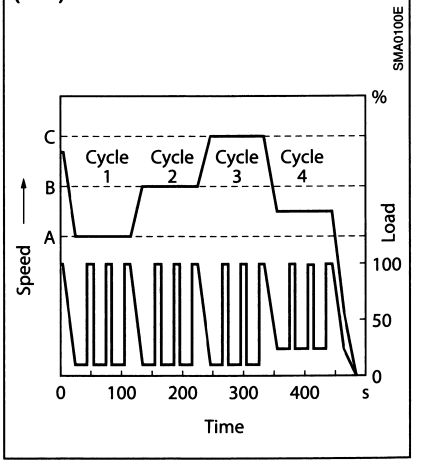


Figure 19: Transient driving cycle ETC (Europe)



European Transient Cycle
As well as Euro III, the ETC (European Transient Cycle, Figure 19) was also introduced to determine gaseous emissions and particulate, and the ELR (European Load Response, Figure 20) test to measure exhaust-gas opacity. Under the Euro III standards, the ETC applies only to commercial vehicles with exhaust-gas treatment (particulate filters, DeNOx catalytic converter); starting with Euro IV (October 2005), it will be obligatory for all vehicles.

Figure 20: European Load Response Test (ELR)

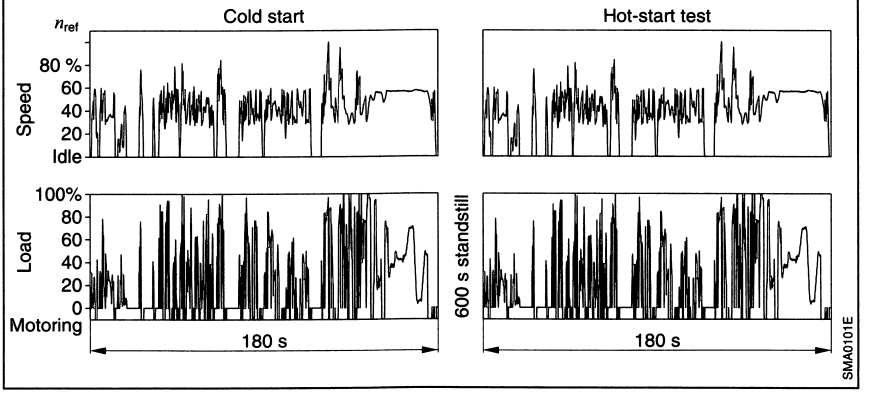


The test cycle is derived from realistic road-driving patterns and is subdivided into three sections: an urban section, an extra-urban section, and an expressway section. The length of the test is 30 minutes, and the periods of time for which engine speeds and torque levels must be maintained are specified in seconds.

World harmonized cycles
Since 2013 world harmonized engine test cycles have had to be applied with the introduction of the Euro VI emission-limit level. The prescribed limits must be equally met both in the WHSC (World Harmonized Stationary Cycle) and in the WHTC (World Harmonized Transient Cycle, Figure 21). A new feature is a WNTE zone (World Harmonized Not To Exceed Zone), as was previously customary only in the USA. The NTE test is conducted in any driving mode within a specified engine-speed/torque range, where an approx. 25 % higher emission limit is permissible.

Compared with the current European tests, the harmonized engine tests are conceived in the direction of lower loads with fewer full-load operating points (Figure 22) and during the transient test clearly more overrun phases. The associated lower exhaust-gas temperatures pose a challenge to active exhaust-gas treatment systems, which must be regularly regenerated.

Figure 21: World Harmonized Transient Cycle (WHTC)



Japan

JE05 test cycle

Pollutant emissions were determined between October 2005 and October 2016 in the transient JE05 test cycle (Figure 23). Similarly to the European transient test for passenger cars, the JE05 test cycle for commercial vehicles comprises an extra-urban part, an urban part, and an expressway part. The test lasts 1,830 seconds and is started when the engine is hot.

Unlike the European and US commercial-vehicle tests, the JE05 test specifies the driving speed instead of the engine speed and engine torque. Because the test is conducted on an engine test bench, the quantities required of engine speed and torque are determined from the specified speeds and from the individual vehicle data using a conversion program. Required quantities include, among others, vehicle weight, tire rolling resistance, transmission ratios, torque curve, and maximum engine speed. With the introduction of the Post PNLT limit stages the harmonized European test cycles WHSC and WHTC are used for certification.

Figure 23: Japanese JE05 test cycle
Average speed: 27.3 km/h.

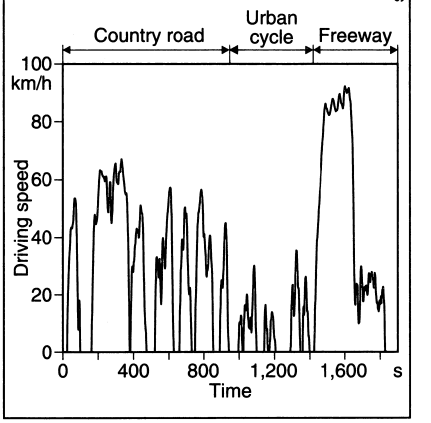
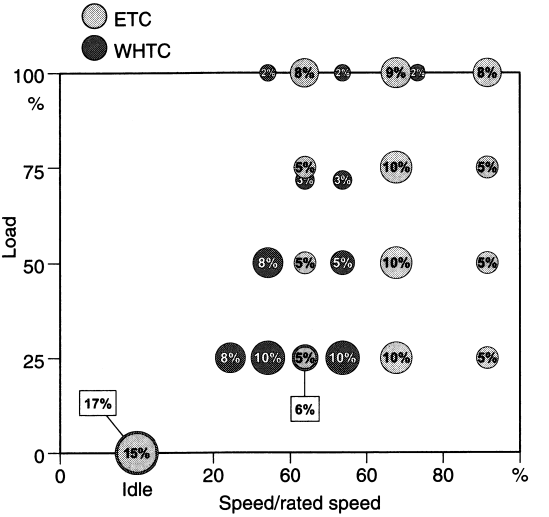


Figure 22: Comparison of ETC and WHTC test cycles



Emission legislation for motorcycles

The USA also led the way in introducing emission regulations for motorcycles. But because motorcycles only constitute a very small percentage of the overall traffic volume and consequently a very small percentage of the overall emissions and immissions, these requirements were for a long time not subject to further development compared with those for passenger cars and commercial vehicles. It was therefore left to the EU and Asian countries such as Japan to drive forward requirements for motorcycles and other L-category vehicles (i.e. two-, three-, and four-wheel vehicles such as mopeds, three-wheel delivery vans, minicars, and quad bikes). The EU regulations were incorporated into ECE Regulations, which are used by countries where two- and three-wheel vehicles are popular such as India, China, and other Southeast Asian countries as the basis for their national legislation. In these countries motorcycles play an important role as a means of transport for persons and goods, with high mileages in both urban and rural areas. Japan, Taiwan, and Thailand have introduced their own regulations for motorcycles.

The targets and instruments such as emission limits, test cycles, test procedures, and monitoring options are basically the same as those for passenger cars and light commercial vehicles (see Emission-control legislation, Overview).

Only the requirements for motorcycles (category L3/L4) with gasoline engines are considered here, as these are the focus of development at Bosch.

US legislation

Exhaust-gas emissions

The EPA introduced the first requirements for motorcycles in 1978. California had between 1988 and 2003 its own requirements, which were extended and tightened from 2004/2008. The EPA introduced only in 2006/2010 new requirements which are based on the Californian requirements from 2004/2008 and carry over to the entire USA.

Motorcycles are subdivided into on-highway and off-highway, each with their own legislation, and encompass two- and three-wheel vehicles. Only on-highway motorcycles are considered here. These have since 2004 (CARB) and since 2006 (EPA) been classified according to their piston displacement into four classes to which the different durability requirements regarding mileage apply:

- Class I-A: < 50 ccm, 5 years / 6,000 km;
- Class I-B: 50 to 169 ccm, 5 years / 12,000 km;
- Class II: 170 to 279 ccm, 5 years / 18,000 km;
- Class III: ≥ 280 ccm, 5 years / 30,000 km.

The exhaust-gas emissions are measured in the FTP cycle on a chassis dynamometer at a temperature of between 20 °C and 30 °C. The emission limits are correlated with the route ridden during the test and are expressed in grams per km. The limits depend on the class and are not fuel-dependent.

Class I-A / I-B and Class II from 2004 (CARB) and from 2006 (EPA):
- carbon monoxide (CO) 12.0 g/km,
- hydrocarbons (HC) 1.0 g/km.

Class III 2004 – 2007 (CARB) and 2006 – 2009 (EPA Tier 1):
- carbon monoxide (CO) 12.0 g/km,
- hydrocarbons (HC) + nitrogen oxides (NO_x) 1.4 g/km.

Class III from 2008 (CARB) and from 2010 (EPA Tier 2):
- carbon monoxide (CO) 12.0 g/km,
- hydrocarbons (HC) + nitrogen oxides (NO_x) 0.8 g/km.

For Class III the manufacturer can average the aggregate limit value for $HC + NO_x$ over its sold fleet.

A manufacturer can optionally certify Class I and Class II vehicles in accordance with the requirements for Class III as well.

There are no requirements with regard to particulate emissions. Unlike passenger cars, there are also no further exhaust-gas requirements.

Evaporative emissions
CARB introduced requirements for evaporative emissions from 1983. Since 2001 all classes have been governed by a limit value of 2.0 g per test in the SHED test (cf. Exhaust-gas measuring techniques), which differs in terms of procedure and duration from the test for passenger cars. Here, after the vehicle has undergone preconditioning including preparation of the activated carbon canister and with the tank filled to the defined level of 40 %, the actual tests start.

1st test: Tank-breathing losses
A linear temperature increase is conducted over 60 minutes in the SHED. Here a distinction is made between exposed and non-exposed tank and between fuel-liquid and fuel-vapor temperature. The hydrocarbon emissions given off by the vehicle are measured.

2nd test: Hot-soak losses
Following the first test the vehicle is heated in a defined riding cycle and then the increase in HC concentration is measured in the SHED over a period of one hour as the vehicle cools down.

The limit value applies to the sum total of the HC emissions in the two tests.

The EPA does not demand a SHED test, but instead focuses on limiting the permeability for hydrocarbons from fuel tank and lines by means of permeation limits. To this end, the vehicle manufacturer can either carry out a permeation test, use EPA-certified materials (defined permeation factors), or install certified components from suppliers.

Permeation limits:
- Fuel tank ≤ 1.5 g/m^2 and day
- Fuel line ≤ 15 g/m^2 and day

There are not requirements for consumption measurement or greenhouse-gas emissions for motorcycles.

EU/ECE legislation
The basis of the EU emission legislation for motorcycles and other L-category vehicles with two, three and four wheels is the General Type Approval Directive 92/61/EEC [8]. With Directive 97/24/EC [9] limits and exhaust-gas emissions were defined on a European level for the first time and have been constantly updated ever since.

As for passenger cars and light commercial vehicles, so too for L-category vehicles the EU legislation serves as the model for the UN Regulations. These feature all the contents of the EU legislation, i.e. the reflect the EU stages described below. New requirements, which are aiming for worldwide harmonization and acceptance, are elaborated on a UN level. This applies in particular to the Worldwide Motorcycle Test Cycle (WMTC) in the Global Technical Regulation GTR No. 2.

The legislative stages for motorcycles are designated Euro or EU and denote an ever-further tightening of the exhaust-gas limits in the Type I test.
- Euro 1 (from 1999)
- Euro 2 (from 2003)
- Euro 3 (from 2006)
- Euro 4 (from 2016)
- Euro 5 (from 2020)

As for passenger cars, a type approval applies to an EU stage up to the point when the next stage is prescribed with binding effect.

The exhaust-gas emissions are measured in a defined test cycle on a chassis dynamometer in the Type I test at a temperature of between 20 °C and 30 °C. The emission limits are correlated with the route ridden during the test and are expressed in milligrams per km. They are in part fuel-dependent, i.e. different for gasoline and diesel engines.

For motorcycles, the test cycle or its applicable parts and their weighting are determined, depending on parameters like piston displacement and top speed. Here, from Euro 3 the synthetic cycle ECE-R40 is superseded by the more realistic dynamic cycle WMTC.

The EU emission legislation specifies for gasoline-engine vehicles limits for the following pollutants:
- Carbon monoxide (CO)
- Hydrocarbons (HC), from Euro 5 also for non-methane hydrocarbons (NMHC)
- Nitrogen oxides (NO_x)
- Particulate matter (PM) (from Euro 5 for gasoline direct-injection engines)

The emission limits for motorcycles (L3) and the test cycles are shown in Table 4. Nominally the Euro 5 limits match the limits for passenger cars Euro 5b.

To obtain approval for each vehicle model, the manufacturer must prove that the emissions of the limited pollutants do not exceed the respective limits over the service life prescribed by law (as distance traveled in kilometers). From Euro 4 for motorcycles with a top speed of 130 km/h or more the service life is 35,000 km, for a top speed of less than 130 km/h it is 20,000 km.

The type-approval test in the Type I test is conducted with a test vehicle which has been broken in for 1,000 km (from Euro 5 2,500 km or 3,500 km for top speed ≤ or > 130 km/h). Since the emission limits apply to the vehicle at the end of its service life, the measured values are multiplied by deterioration factors (from Euro 5 also additive factors) and then compared with the limit values. For this purpose either legally prescribed deterioration factors are used or the manufacturer determines these specially for this vehicle type in the Type V test (vehicle endurance run or from Euro 5 component aging).

Aside from the primary requirement in the Type I test, there is as a further exhaust-gas requirement the Type II test, in which the CO emissions are determined at idle.

As for passenger cars, there is no Type VI test at –7 °C for motorcycles.

Evaporative emissions

As with passenger cars, motorcycles with gasoline engines are subject to requirements to limit evaporative emissions (see Exhaust-gas measuring techniques) in the Type IV test. Evaporative emissions are determined from Euro 4 in the SHED. The procedure and the duration differ from the test for passenger

Table 4: Emission limits Euro 3 to Euro 5 for Class L3 vehicles with gasoline engines (PM only for direct-injection gasoline engine)

Stage	Classification	Cycle/ weighting	CO mg/km	HC mg/km	NMHC mg/km	NO_x mg/km	PM mg/km
Euro 2	Displacement < 150 cm^3	R40 UDC	5,500	1,200	–	300	–
	Displacement ≥ 150 cm^3	R40 UDC	5,500	1,000	–	300	–
Euro 3	Displacement < 150 cm^3	R40 UDC	2,000	800	–	150	–
	Displacement ≥ 150 cm^3	R40 UDC + EUDC	2,000	300	–	150	–
	v_{max} < 130 km/h	WMTC	2,620	750	–	170	–
	v_{max} ≥ 130 km/h	WMTC	2,620	330	–	220	–
Euro 4	v_{max} < 130 km/h	WMTC 30:70	1,140	380	–	70	–
	v_{max} ≥ 130 km/h	WMTC 25:50:25	1,140	170	–	90	–
Euro 5	v_{max} < 130 km/h	WMTC 30:70	1,000	100	68	60	4.5
	v_{max} ≥ 130 km/h	WMTC 25:50:25	1,000	100	68	60	4.5

cars and correspond to the Californian test (see CARB).

The limit value applies to the sum total of the HC emissions in the two tests (tank-breathing and hot-soak losses):
– 2 g/test for Euro 4 (from 1/2016).
– 1.5 g/test for Euro 5 (from 1/2020).

CO_2 emissions
For motorcycles, the CO_2 emissions must also be measured in the same test as for the Type I test and from them the fuel consumption calculated. There are however no fleet targets as for passenger cars and light commercial vehicles.

Japanese legislation
The permissible emissions are being reduced in stages in Japan (similarly to in the EU).

The Japan-specific legislation with a synthetic cycle was replaced from 2012 on the basis of GTR No. 2 by the WMTC, with limit values equivalent to the previous stage.

The next WMTC stage applies from 10/2016 to new domestic types and from 9/2017 to all types and to imports.

A new requirement added from 10/2017 is the limitation of evaporative emissions. The test procedure corresponds to the Californian SHED test with a limit value of 2.0 g HC per test.

From 10/2017 Japan will also introduce its own J-OBD I regulation. This is based on Euro 4-OBD I and on the general OBD requirements of CARB-OBD II. J-OBD II on the basis of Euro 5 is planned for 10/2020.

For motorcycles there are no requirements with regard to consumption measurement or to CO_2 emissions.

Chinese legislation
China has introduced the EU legislation in stages. The stage China 3 from 7/2008 will be replaced from 7/2018 by China 4, which corresponds in terms of exhaust-gas and evaporative emissions to Euro 4. Euro 4-OBD I on the other hand has not been adopted; instead, a China-specific basic OBD I regulation is being developed.

Indian legislation
India has introduced the EU legislation in stages. The stage Bharat IV from 4/2016 corresponds to Euro 3. The stage Euro 4 is to be skipped and from 4/2020 replaced by Bharat VI, which corresponds in terms of exhaust-gas and evaporative emissions to Euro 5. OBD I and OBD II are to be introduced from 4/2020 and from 4/2023 respectively; the data and contents may still change on account of the review of EU OBD II.

References
[1] 70/220/EEC: Council Directive (of the European Communities) of 20 March 1970 on the approximation of laws of the Member States relating to measures to be taken against air pollution by gases from positive-ignition engines of motor vehicles.
[2] 88/77/EEC: Council Directive of 3 December 1987 on the approximation of the laws of the Member States relating to measures to be taken against the emission of gaseous pollutants from diesel engines for use in vehicles.
[3] Regulation (EC) No. 715/2007 of the European Parliament and of the Council of 20 June 2007 on type approval of motor vehicles with respect to emissions from light passenger and commercial vehicles (Euro 5 and Euro 6) and on access to vehicle repair and maintenance information.
[4] Regulation (EC) No. 595/2009 of the European Parliament and of the Council of 18 June 2009 on type approval of motor vehicles and engines with respect to emissions from heavy commercial vehicles (Euro VI) and on access to vehicle repair and maintenance information, and amending Regulation (EC) No. 715/2007 and Directive 2007/46/EC and repealing Directives 80/1269/EEC, 2005/55/EC and 2005/78/EC.
[5] Commission Regulation (EU) No. 64/2012 of 23 January 2012 amending Regulation (EU) No. 582/2011 implementing and amending Regulation (EC) No. 595/2009 of the European Parlia-

ment and of the Council with respect to emissions from heavy commercial vehicles (Euro VI) (Text with EEA relevance).
[6] Commission Regulation (EU) 2016/1718 of 20 September 2016 amending Regulation (EU) No. 582/2011 with respect to emissions from heavy commercial vehicles as regards the provisions on testing with portable emission measurement systems (PEMS) and the procedure for the testing of the durability of replacement pollution control devices (Text with EEA relevance).
[7] Commission Directive 2010/26/EU of 31 March 2010 amending Directive 97/68/EC of the European Parliament and of the Council on the approximation of the laws of the Member States relating to measures against the emission of gaseous and particulate pollutants from internal combustion engines to be installed in non-road mobile machinery.
[8] Council Directive 92/61/EEC of 30 June 1992 relating to the type-approval of two or three-wheel motor vehicles.
[9] Directive 97/24/EC of the European Parliament and of the Council of 17 June 1997 on certain components and characteristics of two or three-wheel motor vehicles.
[10] K. Reif (Editor): Ottomotor-Management – Bosch Fachinformation Automobil. 4th Edition, Springer Vieweg, 2014.

Exhaust-gas measuring techniques

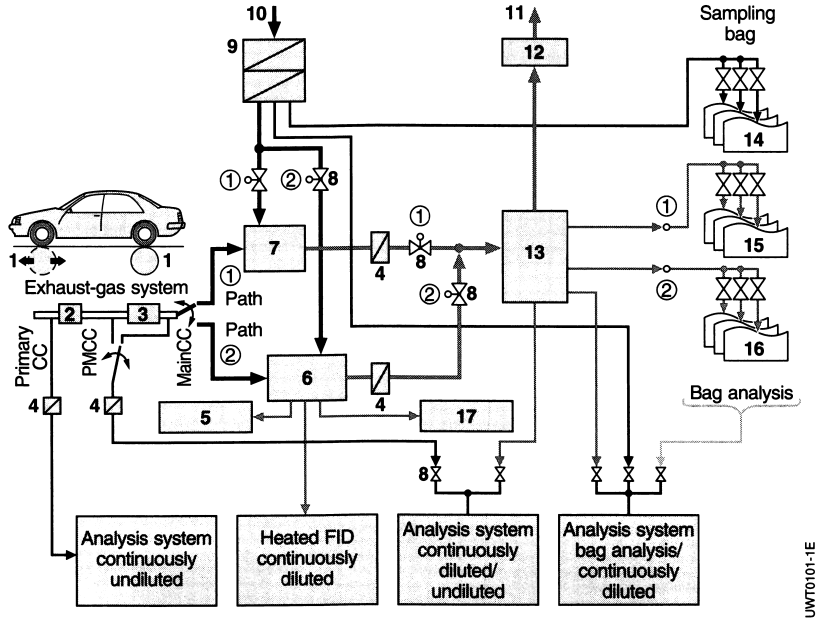
Exhaust-gas testing

Requirements
Exhaust-gas tests on chassis dynamometers are used for the type approval to attain General Certification, as well as to develop engine or other components. They differ from exhaust-gas tests that are conducted in Germany for example in the course of "general and partial inspection for exhaust gas" (§29 StVZO, [1]) using workshop measuring devices. In addition, exhaust-gas tests are carried out on engine test benches, for instance for the type approval of heavy commercial vehicles.

Provided for information purposes only, no liability assumed for completeness!

The exhaust-gas test on chassis dynamometers is performed on vehicles. The methods used are defined to simulate actual vehicle operation on the road as far as possible. Measurement on a chassis dynamometer offers the following advantages here:
- Highly reproducible results, as environmental conditions can be kept constant.

Figure 1: Exhaust-gas test on the chassis dynamometer
1 Roller with dynamometer, 2 Upstream catalytic converter, 3 Underfloor catalytic converter, 4 Filter, 5 Particulate filter, 6 Dilution tunnel, 7 Mix-T, 8 Valve, 9 Dilution-air conditioner, 10 Dilution air, 11 Exhaust-gas/air mixture, 12 Blower, 13 CVS system (Constant Volume Sampling), 14 Dilution-air sample bag, 15 Exhaust-gas sample bag (from Mix-T), 16 Exhaust-gas sample bag (from dilution tunnel), 17 Particulate counter.
① Path for exhaust-gas measurement via Mix-T (without determination of particulate emission),
② Path for exhaust-gas measurement via dilution tunnel (with determination of particulate emission).



- Good comparability of tests, as a defined speed-time profile can be driven independently of traffic flow.
- Stationary setup of the measuring techniques required.

Test setup

General setup

The test vehicle is parked on a chassis dynamometer with its drive wheels on the rollers (Figure 1). This means that the forces acting on the vehicle, i.e. the vehicle's moments of inertia, rolling resistance and aerodynamic drag, must be simulated so that the trip on the test bench reproduces emissions comparable to those obtained during an on-road trip. For this purpose, asynchronous machines, direct-current machines, or even electrodynamic retarders on older test benches, generate a suitable speed-dependent load that acts on the rollers for the vehicle to overcome. More modern machines use electric flywheel simulation to reproduce this inertia. Older test benches use real flywheels of different sizes attached by rapid couplings to the rollers to simulate the vehicle weight. A blower mounted in front of the vehicle provides the necessary engine cooling.

The test-vehicle exhaust pipe is generally a gas-tight attachment to the exhaust-gas collection system – the dilution system is described below. A proportion of the exhaust gas is collected there. At the end of the driving test, the gas is analyzed for gaseous emission-limit components (hydrocarbons, nitrogen oxides and carbon monoxide) and carbon dioxide (to determine fuel consumption).

Following the introduction of the emission-control legislation, particulate emissions were initially limited in diesel-engine vehicles only. In the last few years, legislators have also begun to limit these emissions in vehicles with gasoline engines. To determine particulate emissions, a "dilution tunnel" is used with a high internal flow turbulence (Reynolds number > 40,000). Particulate filters are also used to calculate particulate emission based on load.

In addition, and for development purposes, part of the exhaust gas flow can be extracted continuously from sampling points in the vehicle's exhaust-gas system or dilution system to analyze the pollutant concentrations.

The test cycle is repeated by a driver in the vehicle. The required and current driving speeds are displayed on a driver control-station monitor. In some cases, an automated driving system replaces the driver to increase the reproducibility of test results.

Test setup for diesel-engine vehicles

To determine the pollutant emissions from diesel vehicles, it is necessary to make some changes to the test-bench setup and to the measuring techniques used. The complete sample-taking system, including the exhaust-gas measuring device for hydrocarbons, must be heated to 190 °C. This is to prevent condensation of hydrocarbons which have high boiling points, and to evaporate the hydrocarbons that have already condensed in the diesel exhaust gas.

Dilution system

Purpose of the CVS method

The most commonly used method of collecting the exhaust gases emitted from an engine is the CVS dilution procedure (Constant Volume Sampling). It was introduced for the first time in the USA in 1972 for passenger cars and light-duty commercial vehicles. In the meantime, it has been improved in several stages. The CVS method is used in other countries, such as Japan. It has also been in use in Europe since 1982. It is therefore an exhaust-gas collection method that is recognized throughout the world.

In the CVS method, the exhaust gas is only analyzed at the end of the test. Therefore the condensation of water vapor and the resulting nitrogen-oxide losses and secondary reactions in the collected exhaust gas must be avoided.

Principle of the CVS method

The exhaust gases emitted by the test vehicle are diluted in the Mix-T or in the dilution tunnel with ambient air at a mean ratio of 1:5 to 1:10, and extracted using a special system of pumps in such a

way that the total volumetric flow composed of exhaust gas and dilution air is constant. The admixture of dilution air is therefore dependent on the momentary exhaust-gas volumetric flow. A representative sample is continuously extracted from the diluted exhaust-gas flow and is collected in one or more exhaust-gas sample bags. The sampling volumetric flow is constant during the bag-filling phase. Therefore, the pollutant concentration in a sample bag at the end of the filling process is identical to the mean value of the concentration in the diluted exhaust gas during the bag-filling process.

While the exhaust-gas sample bags are being filled, a sample of the dilution air is taken and collected in one or more air sample bags in order to measure the pollutant concentration in the dilution air.

Filling the sample bags generally corresponds to the phases in which the test cycles are divided (e.g. the ht phase in the FTP 75 test cycle).

The pollutant mass emitted during the test is calculated from the total volume of the diluted exhaust gas and the pollutant concentrations in the exhaust-gas and air-sample bags.

Dilution systems

There are two alternative methods for achieving the constant volumetric flow in the diluted exhaust gas:

- PDP method (Positive Displacement Pump): A rotary-piston blower (Roots blower) is used.
- CFV method (Critical Flow Venturi): A venturi tube and a standard blower are used in the critical state.

Advances in the CVS method

Diluting the exhaust gas causes a reduction in pollutant concentrations as a factor of the dilution. The concentrations of some pollutants (especially hydrocarbon compounds) in the diluted exhaust gas are comparable to the concentrations in the dilution air (or lower) in certain test phases, since pollutant emissions have been significantly reduced in recent years as emission limits have become more stringent. This poses a problem from the measuring-process aspect as the difference in the two values is crucial for the pollutant emissions. A further challenge is presented by the precision of the measuring devices used to analyze the pollutants.

The following measures have been generally implemented in order to counter the problems described above:

- Lowering the dilution; this requires precautions to prevent water from condensing, e.g. by heating sections of the dilution systems, or drying or warming dilution air on vehicles with a gasoline engine.
- Reducing and stabilizing pollutant concentrations in the dilution air, e.g. by using activated charcoal filters.
- Optimizing the measuring devices (including dilution systems), e.g. by appropriately selecting or preconditioning the materials used and system setups or by using modified electronic components.
- Optimizing processes, e.g. by applying special purge procedures.

Bag Mini Diluter

As an alternative to the improvements in CVS technology described above, a new type of dilution system was developed in the USA: the Bag Mini Diluter (BMD). Here, part of the exhaust gas flow is diluted at a constant ratio with dried, heated, pollutant-free zero gas (e.g. cleaned air). During the test, part of this diluted exhaust-gas flow that is proportional to the exhaust-gas volumetric flow is filled in exhaust-gas sample bags and analyzed at the end of the driving test.

In this procedure, dilution is performed with a pollutant-free zero gas free of pollutants and not with air containing pollutants. This has the purpose of avoiding the air-sample bag analysis and the subsequent differential formation of exhaust-gas and air-sample bag concentrations. However, a more complex procedure is required than that for the CVS method, e.g. one requirement is to determine the (undiluted) exhaust-gas volumetric flow and the proportional sample-bag filling.

Exhaust-gas measuring devices

In vehicles with gasoline engines, the emissions of limited gaseous pollutants are calculated from the concentrations in exhaust-gas and air sample bags. Emission-control legislation defines globally standard test procedures for this purpose (Table 1).

Essentially, the same devices are used to measure the concentrations of gaseous pollutants in the exhaust gas of gasoline-engine vehicles as for diesel-engine vehicles. There is however a difference in how the hydrocarbon emissions (HC) are determined. This is done not in the exhaust-gas sample bag, but through continuous analysis of a partial flow of diluted exhaust gas and integration of the measured concentration via the driving test. The reason for this is that the hydrocarbons (which have a high boiling point) condense in the (non-heated) exhaust-gas sample bag.

Table 1: Discontinuous measuring methods

Components	Procedure
CO, CO_2	Non-dispersive infrared analyzer (NDIR)
Nitrogen oxides (NO_x)	Chemiluminescence detector (CLD).
Total hydrocarbon (THC)	Flame ionization detector (FID)
CH_4	Combined design of gas chromatographic procedure and flame ionization detector (GC FID)
CH_3OH, CH_2O	Combined design of impinger or cartridge process and chromatographic analysis techniques; mandatory in the USA when certain fuels are used
Particulates	1.) Gravimetric method: Weighing of particulate filters before and after the test drive 2.) Particulate counting

For development purposes, many test benches also include the continuous measurement of pollutant concentrations in the vehicle exhaust-gas system or the dilution system. The reason is to capture data for the components under control, as well as for other components not subject to legislation. Other test procedures than those listed in Table 1 are required for this, e.g.:

- Paramagnetic method (to measure O_2 concentration).
- Cutter FID: a combination of flame ionization detector and absorber for non-methane hydrocarbons (to measure the CH_4 concentration).
- Mass spectroscopy (multi-component analyzer).
- FTIR spectroscopy (Fourier Transform Infrared, multi-component analyzer).
- IR laser spectroscopy (multi-component analyzer).

A description of the main measuring devices is given below.

Test-bench measuring techniques

NDIR analyzer

The NDIR (Non-Dispersive Infrared) analyzer utilizes the property of certain gases to absorb infrared radiation within a narrow wavelength range. Absorbed radiation is converted to vibration or rotation energy by the absorbing molecules. In turn, this energy can be measured as heat. The phenomenon described occurs in molecules that are formed from atoms of at least two different elements, e.g. CO, CO_2, C_6H_{14} or SO_2.

There are a number of variants of NDIR analyzers; the main component parts are a source of infrared light (Figure 2), an absorption cell (cuvette) through which the test gas is routed, a reference cell generally positioned in parallel (filled with inert gas, e.g. N_2), a rotating chopper, and a detector. The detector comprises two chambers connected by a membrane and containing samples of the gas components under analysis. Radiation from the reference cell is absorbed in one

chamber and radiation from the cuvette in the other. Radiation from the cuvette may have already been reduced by absorption in the test gas. The difference in radiant energy causes a flow movement that is measured by a flow sensor or a pressure sensor. The rotating chopper interrupts the infrared radiation in cycles, causing the flow movement to change direction and therefore a modulation of the sensor signal.

NDIR analyzers possess strong cross sensitivity to water vapor in the test gas since H_2O molecules absorb a wide range of infrared radiation wavelengths. This is the reason why NDIR analyzers are positioned downstream of a test-gas treatment system (e.g. a gas cooler) to dry the exhaust gas when they are used to make measurements on undiluted exhaust gas.

Chemiluminescence detector (CLD)
In a reaction chamber, the test gas is mixed with ozone that is produced from oxygen in a high-voltage discharge (Figure 3). The nitrogen monoxide content in the test gas oxidizes to nitrogen dioxide in this environment; some of the molecules produced are in a state of excitation. When these molecules return to their basic state, energy is released in the form of light (chemiluminescence). A detector (e.g. photomultiplier) measures the emitted luminous energy; under specific conditions, it is proportional to the nitrogen-monoxide concentration (NO) in the test gas.

It is a requirement to measure the NO and NO_2 molecules as the legislation regularizes the emission of the total nitrogen oxides. However, since the test principle of the chemiluminescence detector is limited to measuring the NO concentration, the test gas is channeled through a converter that reduces the nitrogen dioxide to nitrogen monoxide.

Flame Ionization Detector (FID)
The test gas is burned in a hydrogen flame (Figure 4), where carbon radicals are formed and some of these radicals are temporarily ionized. The radicals are discharged at a collector electrode. The current produced is measured and is proportional to the number of carbon atoms in the test gas.

Figure 2: Test chamber acc. to the NDIR method
1 Gas outlet, 2 Absorption cell,
3 Test-gas inlet, 4 Optical filter,
5 Infrared light source, 6 Infrared radiation,
7 Reference cell, 8 Rotating chopper,
9 Detector.

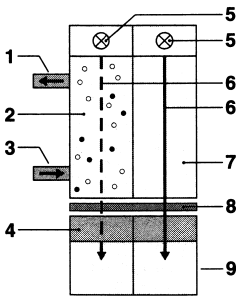
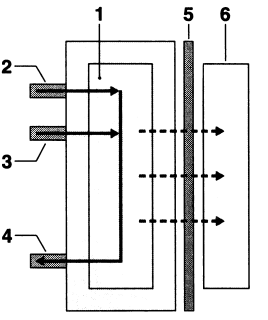


Figure 3: Design of chemiluminescence detector
1 Reaction chamber, 2 Ozone inlet,
3 Test-gas inlet, 4 Gas outlet,
5 Filter, 6 Detector.

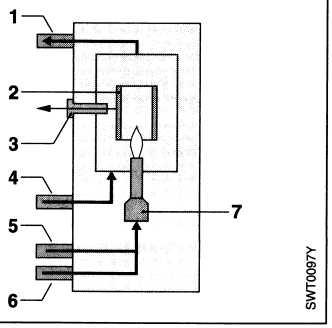


GC FID and Cutter FID
There are two generally common methods to measure the methane concentration (CH_4) in the test gas. Each method consists of the combination of a CH_4-separating element and a flame ionization detector. In these methods, either a gas-chromatography column (GC FID) or a heated catalytic converter oxidizing the non-CH_4 hydrocarbons (cutter FID) are used to separate the methane.

Unlike the cutter FID, the GC FID can only determine the CH_4 concentration discontinuously (typical interval between two measurements: 30 to 45 s).

Paramagnetic detector (PMD)
There are different constructions of paramagnetic detectors (dependent on the manufacturer). The constructions are based on the phenomenon that forces with paramagnetic properties (such as oxygen) act on molecules in inhomogeneous magnetic fields. These forces cause the molecules to move. The movement is sensed by a special detector and is proportional to the concentration of molecules in the test gas.

Figure 4: Design of flame ionization detector
1 Gas outlet, 2 Collector electrode, 3 Amplifier output, 4 Combustion air, 5 Test-gas inlet, 6 Combustion gas (H_2/He), 7 Burner.



Measuring particulate emission
In addition to gaseous pollutants, solid particulates are also measured, as they are also pollutants subject to legislation. Currently, the gravimetric process is the process specified by law to measure particulate emissions.

Gravimetric process (particulate-filter process)
Part of the diluted exhaust gas is sampled from the dilution tunnel (CVS method) during the driving test and then channeled through particulate filters. The particulate loading is calculated from the weight of the particulate filters before and after the test. The particulate emission during the driving test is then calculated from the load, the total volume of the diluted exhaust gas, and the partial volume channeled through the particulate filters.

The gravimetric process has the following disadvantages:
- Relatively high detection limit, only reducible to a limited extent by using complex instrumentation (e.g. to optimize the tunnel geometry).
- It is not possible to measure particulate emissions continuously.
- The process is complex as particulate filters have to be conditioned in order to minimize environmental influences.
- It is not possible to select the chemical composition of particulates or particulate size.

Particulate counting
Because of the disadvantages mentioned and the ongoing reduction of the limit values, increasingly the number of emitted particulates will be limited, as well as the particulate emission (particulate mass per distance traveled).

The "Condensation Particulate Counter" (CPC) has been earmarked as the measuring device for calculating the number of particulates in compliance with legislation (particulate counting). In this counter, a small partial flow of the diluted exhaust gas (aerosol) is mixed with saturated butanol vapor. The process of the butanol condensing on the solid particulates causes the particulates

to increase dramatically in size so that it is possible to calculate the number of particulates in the aerosol with the aid of scattered-light measurement.

The number of particulates in the diluted exhaust gas is continuously calculated; integration of the measured values produces the number of particulates over the driving test.

Determination of particulate size distribution
There is increasing interest in acquiring knowledge of the size distribution of particulates in the exhaust gas of a vehicle. Examples of devices that supply this information are the Scanning Mobility Particle Sizer (SMPS), Electrical Low Pressure Impactor (ELPI), and Differential Mobility Spectrometer (DMS).

Portable emission measuring devices
A PEMS measuring device (Portable Emission Measurement Systems) consists primarily of three components to ensure a complete analysis of the vehicle emissions.

Gas PEMS
Gas-analysis apparatus is used to measure the pollutant emissions in the exhaust-gas flow. This consists of an NDUV analyzer (Non-Dispersive Ultraviolet analyzer) to determine NO and NO_2, an NDIR analyzer (see Test-bench measuring techniques) to measure CO and CO_2, and an O_2 sensor.

The NDUV analyzer is a multi-component UV photometer. The measuring principle is based here essentially on the characteristic attribute of NO and NO_2 to absorb luminous radiation in a wavelength range of 200 to 500 nm.

PN PEMS (Particle Number)
Two different measuring principles are used to measure the particulates. Firstly, portable "Condensation Particle Counters" (CPCs) are used which through condensation of the particulates, for example with n-butanol vapor make it possible to detect nanoscale particulates optically (size range: 23 to 700 nm) (see Test-bench measuring techniques, particulate counting).

Secondly, portable "Diffusion Chargers" (DCs) are used to count particulates. The measuring principle is based on the electrostatic precipitation of charged particulates. The particulates are charged by collisions with ions which are created by a gas discharge in a charging chamber in the measuring device. Precipitation of the charged particulates creates a potential difference between the metallic precipitation surface and the surrounding area. It is then possible by means of numerical methods to use this potential difference to determine theoretically the particulate number.

EFM (Exhaust Flow Meter)
The exhaust-gas volumetric flow is measured with an exhaust flow meter (EFM). Here the pressure loss is measured at a defined flow resistance inside the test pipe. The differential pressure can then be used to calculate the volume of exhaust gas.

The EFM is a test pipe mounted on the exhaust tailpipe which contains both pressure and temperature sensors. The sampling point for the gas-analysis apparatus and the sampling point for measuring the particulate number are also integrated on this adapter.

Testing commercial vehicles
In the USA the transient test method has been prescribed since model year 1987 for testing the emissions from diesel engines in heavy commercial vehicles over 8,500 lbs (EPA) and 14,000 lbs (CARB) gross vehicle weight. It is performed on dynamic engine test benches. In Europe this test method was introduced with Euro III for engines with active exhaust-gas-treatment (particulate filter or selective catalytic reduction, SCR) and with Euro IV for all engines in heavy commercial vehicles over 3.5 t gross vehicle weight. Since Euro V the test on

the engine test bench is dependent no longer on the gross vehicle weight but on the reference weight (tare weight plus 100 kg).

The transient test method also uses the CVS method. However, the size of the engines demands a test setup with a substantially higher throughput in order to keep to the same dilution ratios as for cars and light-duty commercial vehicles. Double dilution (through a secondary tunnel) approved by legislators helps to limit the increased complexity of instrumentation. The diluted exhaust-gas volumetric flow can be set alternatively using a Roots blower or critical-flow venturi nozzles.

Another possibility would be to calculate the particulate emissions with a partial-flow dilution system, provided that the remaining pollutants in the undiluted exhaust gas are measured.

With Euro VI (2013) limit values for the particulate number are binding on all new model types. The value is $8 \cdot 10^{11}$ particles in steady-state and $6 \cdot 10^{11}$ particles per kWh in transient operation.

Diesel smoke-emission test

Methods

Separate legislation for testing the smoke emissions of diesel-engine vehicles came into force long before the introduction of legislation for testing gaseous pollutants. All existing smoke tests are coupled closely with the measuring devices used. One measure of smoke emission (soot, particulates) is the smoke number.

Two methods are essentially customary for measuring this value. In the absorption method (opacity measurement), the opacity of the exhaust gas is indicated by the degree to which it blocks the passage of a beam of light shining through it (Figure 5). In the filter method (measurement of reflected light), a specified quantity of exhaust gas is routed through a filter element. The degree of filter discoloration provides an indication of the amount of soot contained in the exhaust gas (smoke tester, Figure 7).

Measurement of diesel smoke emissions is relevant only if the engine is under load, since it is only when the engine is operated under load that emission of significant levels of particulates occur. Two different test procedures are also commonly used here. Firstly, measurements under full load, e.g. on a chassis dynamometer, or on a defined test circuit. Secondly, measurements under unrestricted acceleration using a defined throttle burst under the load of the engine flywheel (Figure 6).

As the results of testing for diesel smoke emissions vary according to both test procedure and type of load, they are not generally suitable for direct mutual comparisons.

Opacimeter (absorption method)

Exhaust-gas opacity denotes the attenuation of light by absorption, diffraction, scatter and reflection on the particulates contained in the exhaust gas.

A transmitter and a photodetector are mounted on the exhaust pipe for measurement of the full flow with the opacimeter. In partial-gas devices, the exhaust gas is routed through an exhaust-gas

sample probe and heated lines by a pump into the measuring chamber (Figure 5). Using a longer measuring chamber improves the device's detection limit.

During the unrestricted acceleration part of the exhaust gas is delivered to the measuring chamber. A light beam passes through the measuring chamber now filled with exhaust gas. The attenuation of lights is determined photoelectrically and indicated as opacity T in % or as an absorption coefficient k in m^{-1} (Figure 6). A precisely defined measuring-chamber length and keeping the optical window clear of soot (by air curtains, i.e. tangential air flows) are basic requirements for high levels of accuracy and reproducibility of the measurements.

During testing under load, measurement and display are a continuous process. The opacimeter automatically determines the maximum value and calculates the mean from several gas pulses.

Smoke tester (filter method)

During soot measurement according to Bosch a defined exhaust-gas volume (330 cm^3) is drawn by means of a hand pump through a white filter paper. The blackening of the paper is determined by means of a reflex photometer. The degree of paper blackening is divided

Figure 6: Gas-pulse measurement with the opacimeter

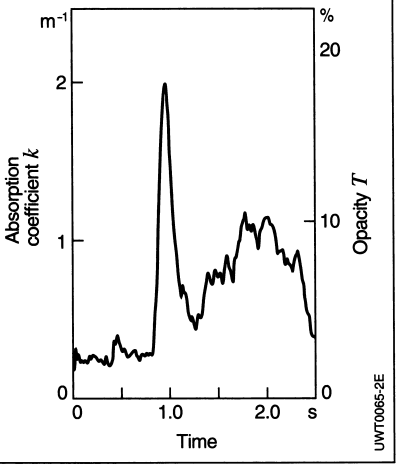
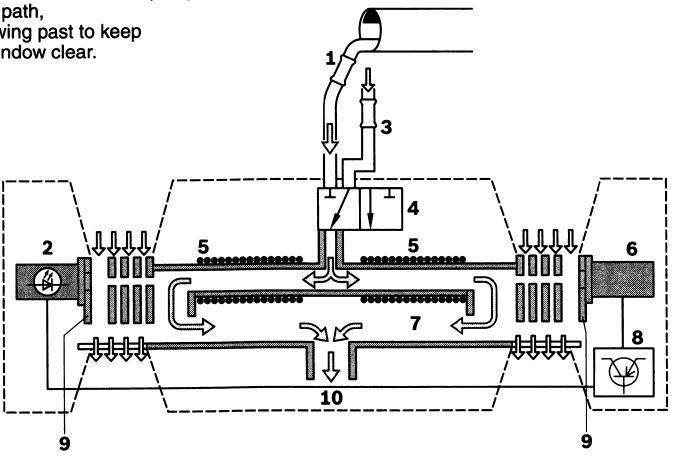


Figure 5: Opacimeter (absorption method)
1 Sample probe, 2 Light source, 3 Purge air for calibration,
4 Calibrating valve, 5 Heater, 6 Receiver, 7 Measuring chamber,
8 Electronic analyzer and display, 9 Optical window,
10 Measuring-chamber outlet, to pump.
➧ Exhaust-gas path,
⇨ Purge air flowing past to keep the optical window clear.

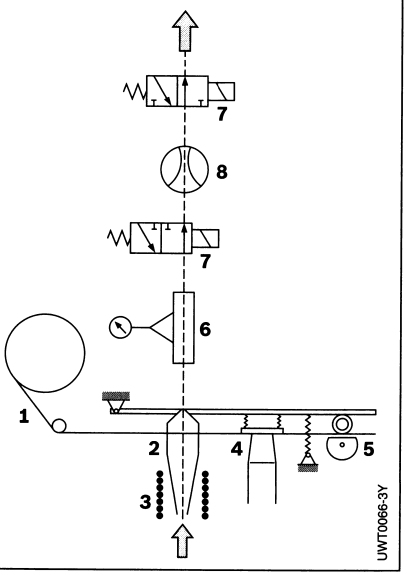


into numbers ranging between 0 and 10 (smoke number according to Bosch, Bosch number), where 0 denotes white, unused paper and 10 denotes fully blackened measuring points.
An empirical correlation can be used to convert to soot-mass concentration in mg/m^3.

The method has been developed over decades. Thus, the hand pump, for example, was replaced by a continuously operating pump, the dead volume between the sample probe and the filter paper is taken into consideration, and the devices are heated to prevent condensation (Figure 7).

The filter smoke number (FSN) is defined in ISO 10054 [2]. It corresponds to the Bosch number for an effective suction length of 405 mm at a temperature of 298 K and a pressure of 1 bar.

Figure 7: Smoke tester (filter method)
1 Filter paper, 2 Gas passage,
3 Heater, 4 Reflex photometer,
5 Paper transport, 6 Volume measurement,
7 Purge-air changeover valves,
8 Pump.



Evaporative-emissions test

Independently of the combustion pollutants produced in the engine, a gasoline-engine vehicle emits additional quantities of hydrocarbons (HC) through evaporation of fuel from the fuel tank and fuel system. The quantity evaporated is dependent on the design of the vehicle and on the fuel temperature. An activated carbon canister, in which the fuel vapors are stored, is generally used as a limiting measure. Since it only has a limited absorption capacity, the activated carbon canister must be regenerated at certain intervals. This is done by purging with fresh air while driving. The air/fuel mixture is delivered to the induction tract and combusted in the engine (see Evaporative-emissions control system). Many countries (e.g. the USA and Europe) have regulations which limit these evaporative losses.

Test methods

These evaporative emissions are usually quantified with the aid of a hermetically sealed climate chamber, the SHED (Sealed Housing for Evaporative emissions Determination). For the test, HC concentrations are measured at the beginning and the end of the test with a flame ionization detector (FID), with the difference representing the evaporative losses.

The evaporative losses must – depending on the country – be measured in some or in all of the operating states below and satisfy emission limits:

- Evaporation that emerges from the fuel system when the vehicle is parked with the engine warm following operation: hot-soak test (EU, USA, and others).
- Evaporation emerging from the fuel system as a result of temperature changes in the course of the day: tank-breathing test or diurnal test (EU, USA, and others).
- Evaporative emissions while driving, e.g. due to permeation: running-loss test (USA only).

Test procedures
Evaporations are measured in several phases during a detailed prescribed test procedure. After the vehicle has undergone preconditioning including preparation of the activated charcoal canister ad with the tank filled to the defined level of 40 %, testing starts.

First test: Hot-soak losses (hot soak)
To measure the evaporative emissions in this test phase, the vehicle is heated up before the test by running through the test cycle valid for the country concerned. The engine is then turned off when the vehicle is in the climate chamber. The increase in HC concentration is measured for a period of one hour as the vehicle cools.

The vehicle's windows and trunk lid must remain open during the test. This allows evaporative losses from the vehicle interior to be included in measurements.

Second test: Tank-breathing losses
For this test, a typical temperature profile for a warm summer day (maximum temperature for EU: 35 °C; for EPA: 35.5 °C; for CARB: 40.6 °C) is simulated inside the hermetically sealed climate chamber. The hydrocarbons emitted by the vehicle under these conditions are then collected.

- In the USA, both a 2-day diurnal test (48 hours; in-use test) and a 3-day diurnal test (72 hours; certification) must be carried out (the highest value over the course of the day is used in each case).
- EU legislation from Euro 3 provides for a 24-hour test. From 2019 a revised MNEDC-based evaporative-emissions test will come into force, with changes in the conditioning procedure (less occasion for purging) and using an aged activated carbon filter, a new permeation test for the tank system, and extending the tank-breathing test to 48 hours (it counts the sum total of the HC values of the two 24-hour cycles). The limit value is still 2 g HC/test.

On the basis of the new EU test a WLTC-based test was devised on a UN level in 2016/17 as GTR 19; this is intended to replace the MNEDC test after a transition phase.

Third test: Running losses
The running-loss test is conducted prior to the hot-soak test. It is used to assess the hydrocarbon emissions generated during vehicle operation in the prescribed test cycles (1 x FTP 72, 2 x NYCC, 1 x FTP 72; see USA test cycles).

Limits
EU legislation
The total of the measurement results from the first and second tests produces the evaporative losses. This total must lie below the current limit value of 2 g of evaporated hydrocarbons for all the measurements. The result of the permeation test will be added from 2019.

USA
In the USA (legislation according to CARB LEV II and EPA Tier 2), the evaporative losses monitored in the running-loss test must remain below 0.05 g per mile. The limit values for the hot-soak losses and the tank-breathing losses are defined as follows:

- 2-day diurnal + hot soak:
  1.2 g (EPA) / 0.65 g (CARB),
- 3-day diurnal + hot soak:
  0.95 g (EPA) / 0.50 g (CARB).

These limits must be adhered to over 120,000 miles (EPA) or 150,000 miles (CARB). They have been introduced in stages since model year 2004 and have applied in full since model year 2007 (EPA) and 2006 (CARB). With model year 2009, EPA alternatively permits certification in accordance with CARB limits and CARB regulations (harmonization).

For PC and LDT1, which are certified in accordance with the CARB ZEV legislation, a lower limit value in the SHED test of 0.350 g HC per test and additionally the "zero evaporative emissions" requirement apply. Practically no emissions of

fuel (limit value: 0.054 g HC/test) are meant. To this end, the above-described "3-day diurnal + hot soak" is carried out in the "rig test", i.e. a setup of tank, fuel lines, activated carbon canister, and engine by arrangement between CARB and the vehicle manufacturer.

With LEV III and Tier 3 the requirements for limiting evaporative emissions of the ZEV legislation is extended to all vehicles. The manufacturers can either adhere to the above-described ZEV limit values or alternatively the combination of slightly stricter limit values in the SHED test (for PC/LDT1 0.300 g HC/test) and the new BETP test ("Bleed Emissions Test Procedure"), with which the tightness and the purging performance only of tank, fuel lines and carbon canister but without engine are tested (limit value for PC/LDT1 0.020 g HC per test). The durability requirement is 150,000 miles.

China

China has adopted up to and including Stage 5 the EU requirements for limiting evaporative emissions. From China 6a (7/2020) China will be introducing its own legislation which combines the US procedure with WLTC elements. The limit value for the sum total of hot-soak test and 2-day diurnal test (max. temperature 38 °C, counting the higher 24-hour HC value) is 0.7 g HC.

Refueling emissions

Refueling test

In the refueling test the HC emissions when refueling are measured in order to monitor evaporation of the fuel vapors displaced during refueling (limit value: 0.053 g HC per liter of pumped fuel). In the USA, this test applies to both CARB and EPA.

From China 6a (7/2020) China will be introducing its own legislation which combines the US procedure with WLTC elements. The limit value is 0.05 g HC/test.

Spitback test

In the spitback test, the quantity of fuel splashed out during each refueling operation is measured. The tank must be refueled to at least 85 % of its total volume. This test is only carried out if the refueling test was failed (limit value: 1 g HC per test).

References

[1] § 29 StVZO: Untersuchung der Kraftfahrzeuge und Anhänger (Inspection of motor vehicles and trailers).

[2] ISO 10054: Internal combustion compression-ignition engines – Measurement apparatus for smoke from engines operating under steady-state conditions – Filter-type smokemeter (1998).

Diagnostics

The rise in the sheer amount of electronics in the vehicle, the use of software to control the vehicle, and the increased complexity of modern electronic systems place high demands on the diagnostic concept, monitoring in driving mode (on-board diagnostics), and workshop diagnostics.

As emission-control legislation becomes more and more stringent and continuous monitoring in driving mode is now called for, lawmakers have now acknowledged on-board diagnostics as an aid to monitoring exhaust-gas emissions, and have produced manufacturer-independent standardization. This additional system is termed the OBD system (On-Board Diagnostic system). Diagnosis of engine management systems is thus particularly important.

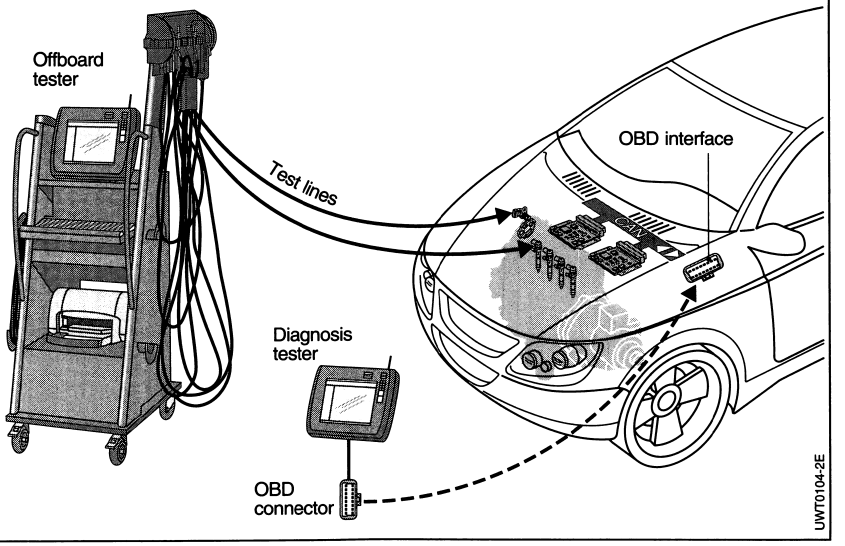
Provided for information purposes only, no liability assumed for completeness!

Monitoring in driving mode

The diagnosis integrated in the control unit is a basic feature of electronic engine management systems. Besides a self-test of the control unit, input and output signals, and control-unit intercommunication are monitored.

Monitoring algorithms check input and output signals during vehicle operation, and check the entire system and all the relevant functions for malfunctions and disturbances. Any errors or faults detected are stored in the control-unit fault memory. When the vehicle is serviced in the dealer's workshop, the stored information is exported over a serial interface. This allows troubleshooting and repairs to be carried out quickly and reliably (Figure 1).

Figure 1: Diagnostic system consisting of a diagnostic tester for reading out OBD data and an external offboard tester for further selective troubleshooting and fault localization



Monitoring input signals
The sensors, connectors and connecting lines (in the signal path) to the control unit are monitored by using the evaluated input signals. With this monitoring strategy, it is possible to detect sensor faults, short circuits to battery voltage U_B and to ground, and open circuits. The following procedures are used for this purpose:
- Monitoring sensor supply voltage (if applicable)
- Examining measures recorded for permissible value ranges (e.g. 0.5 to 4.5 V)
- Plausibility check of different physical signals (e.g. comparison of crankshaft and camshaft speeds)
- Plausibility check of a physical variable which is detected redundantly with different sensors (e.g. pedal-travel sensor)

Monitoring output signals
Actuators triggered by a control unit via output stages are monitored. The monitoring functions detect open circuits and short circuits in addition to actuator faults. The following procedures are used for this purpose: On the one hand, the electric circuit of an output signal is monitored by the output stage; the circuit is monitored for short-circuits to battery voltage U_B, to ground, and for open circuit. On the other hand, impacts on the system by the actuator are detected directly or indirectly by a function or plausibility monitor. System actuators, e.g. the exhaust-gas recirculation valve or the throttle valve, are monitored indirectly via closed-control loops (e.g. for continuous control variance), and also partly by means of position sensors (e.g. position of throttle valve).

Monitoring of internal ECU functions
Monitoring functions are implemented in control-unit hardware (e.g. "intelligent" output-stage modules) and software to ensure that the control unit functions correctly at all times. The monitoring functions check each of the control-unit components (e.g. microcontroller, flash EPROM, RAM). Many tests are performed immediately after switch-on. Further monitoring functions are repeated at regular intervals during normal operation so that component failure can also be detected during operation. Test procedures that require a high amount of computing capacity, or which cannot be performed during vehicle operation for other reasons, are carried out in the after-run after "engine off". This avoids impacting the other functions. An example of such a function is the checksum check of the flash EPROM.

Monitoring ECU communication
Communication with the other control units normally takes place over the CAN bus. Control mechanisms for error detection are incorporated in the CAN protocol so that transmission errors can be detected in the CAN chip already. The control unit also runs a variety of other tests. Since the majority of CAN messages are sent at regular intervals by the individual control units, the failure of a CAN controller in a control unit is detectable by checking these regular intervals. If redundant information is stored in the control unit, all input signals are checked against this information.

Error/fault handling
Error/fault detection
A signal path (e.g. sensor with plug connector and connecting line) is categorized as totally defective if a fault occurs over a specific length of time. The system will continue to use the last valid value until the defect is categorized. When the defect is categorized, a standby function is triggered (e.g. engine-temperature substitute value $T = 90\,°C$).

A healing or "restored-signal recognition" feature is available during driving mode for most errors. The signal path must be detected as intact for a specific period of time for this purpose.

Error/fault storage
All faults are stored as a fault code in the non-volatile area of the data memory. The fault code also describes the fault type (e.g. short circuit, open circuit, plausibility, value range exceeded). Each fault-code input is accompanied by additional information, e.g. the operating con-

ditions (freeze frame) at the time of fault occurrence (e.g. engine speed, engine temperature).

Limp-home functions
If a fault is detected, limp-home strategies can be triggered in addition to substitute values (limp home, e.g. engine output power or speed limited). These measures are used to maintain driving safety, prevent consequential damage (e.g. catalytic converter overheating), and minimize exhaust-gas emissions.

On-board diagnostics

The engine system and components must be continuously monitored in driving mode so that compliance with the emission limits required by law can be achieved in everyday use. Therefore, starting in California, regulations were adopted to monitor exhaust-gas-related systems and components. This has standardized and expanded manufacturer-specific on-board diagnostics (OBD) with respect to the monitoring of emission-related components and systems [6].

OBD I (CARB)
In 1988 the first stage of CARB legislation (California Air Resources Board) came into force in California with OBD I. This first OBD stage requires the monitoring of emission-related electrical components (short-circuits, line breaks) and storage of the faults in the control-unit fault memory as well as a malfunction indicator lamp (MIL) that alerts the driver to detected faults. On-board means (e.g. flashing code on a diagnosis lamp) must also be in place to provide a readout of which component has malfunctioned.

OBD II (CARB)
In 1994 the second stage of diagnosis legislation was introduced in California with OBD II. OBD II became mandatory for diesel-engine cars with effect from 1996. In addition to the scope of OBD I, system functionality was now monitored (e.g. plausibility check of sensor signals).

OBD II stipulates that all emission-related systems and components must be monitored if they cause an increase in toxic exhaust-gas emissions (and thus the OBD threshold values to be exceeded) in the event of a malfunction. In addition, all components used to monitor emission-related components or which affect the diagnosis result must be monitored.

Normally, the diagnostic functions for all components and systems under surveillance must run at least once during the exhaust-gas test cycle (e.g. FTP 75, Federal Test Procedure).

The OBD II legislation furthermore prescribes standardization of the fault-memory information and access to the information (connector, communication) compliant with ISO 15031 [1] and the corresponding SAE standards (Society of Automotive Engineers), e.g. SAE J1979 [2] and SAE J1939 [3]. This permits a fault-memory readout over standardized, commercially available testers (scan tools).

OBD II expansions

From model year 2004

The law has been revised several times since OBD II was introduced. The legal requirements are generally revised by the authorities every two years ("biennial review"). Since model year 2004 it has been necessary, in addition to meeting tighter and additional functional requirements, to check the diagnostic frequency (from model year 2005) in everyday operation (In-use Monitor Performance Ratio, IUMPR).

From model year 2007 through 2013 for gasoline passenger cars

New requirements for gasoline engines are essentially the diagnosis of cylinder-individual mixture trimming (air/fuel imbalance), extended requirements with regard to diagnosis of the cold-start strategy, and permanent error/fault storage, which also applies to diesel systems.

From model year 2007 through 2013 for diesel vehicles

For diesel passenger cars and light commercial vehicles the tightened OBD emission limits are classified into three stages (up to model year 2009, model years 2010 through 2012, from model year 2013). Significantly extended functions are additionally required for the fuel-injection system, the air system, and the exhaust-gas treatment system. Thus, for example, the fuel-injection system is required to monitor the injected fuel quantity and the injection timing. The air system is required, for example, to monitor boost-pressure control and additionally the dynamic response of exhaust-gas recirculation control and boost-pressure control. In the exhaust-gas treatment system, new monitoring functions are required of the oxidation-type catalytic converter, the particulate filter, the NO_x storage catalytic converter, and the SCR dosing system (Selective Catalytic Reduction) with SCR catalytic converter. Thus, to be monitored are, for example, the regeneration frequency in the particulate filter and the dosing quantity of the NO_x reducing agent in the SCR dosing system.

One of the new requirements for diesel systems since 2009 is that, as well as the regulators, controlled functions – where they are relevant to exhaust gas – are to be monitored as well. Likewise, the monitoring of cold-start functions is subject to extended requirements.

From model year 2014/2015

Enhanced requirements have already been formulated for model year 2015 for individual components in diesel passenger cars, light commercial vehicles, and general commercial vehicles. These relate to the monitoring of the oxidation catalytic converter for "feedgas" (ratio between NO and NO_2 to operate the SCR catalytic converter), to the monitoring of the coated particulate filter for NMHC conversion (non-methane hydrocarbons), and to the monitoring of the injection system for quantity-encoded injectors. Likewise, in the course of the revision of the LEV III emission legislation some requirements pertaining to hybrid vehicles were defined more precisely with an impact on the IUMPR calculation.

From model year 2017/2023

The last revision of the OBD legislation saw among others an adaptation of the OBD threshold values to the LEV III emission legislation. The OBD II threshold value for NO_x and NMHC will be defined from LEV III as a quantity (NO_x + NMHC). The OBD threshold values as multiples of the emission limit (multipliers) were adapted in stages to the new emission categories (ULEV 50, ULEV 70, SULEV 20). For the first time an OBD threshold value was also de-

Table 1: OBD threshold values

	Gasoline passenger cars	Diesel passenger cars	Diesel commercial vehicles
CARB	– Dependent on the emission category and diagnostic requirement between 1.5 and 2.5 times the emission limit. – PM OBD limit: absolute 17.5 mg/mile.	– Dependent on the emission category and diagnostic requirement between 1.5 and 2.5 times the emission limit. – PM OBD limit: absolute 17.5 mg/mile – But from 2007 to 2013 introduction of stricter limits in three stages: e.g. for particulate filter 2007 – 2009 5 x limit 2010 – 2012 4 x limit from 2013 1.75 x limit	2010–2012: CO: 2.5 x limit NMHC: 2.5 x limit NO_x: +0.4/0.6 g/bhp-hr [2] PM: +0.06/0.07 g/bhp-hr From 2013: CO: 2.0 x limit NMHC: 2.0 x limit NO_x: +0.2/0.4 g/bhp-hr [2] PM: +0.02/0.03 g/bhp-hr Transition phase for some monitors up to 2016.
EPA (US Federal)	see CARB CARB certificates with corresponding limits are recognized by EPA	see CARB CARB certificates with corresponding limits are recognized by EPA	2010–2012: CO: 2.5 x limit NMHC: 2.5 x limit NO_x: +0.6/0.8 g/bhp-hr [2] PM: +0.04/0.05 g/bhp-hr From 2013: CO: 2.0 x limit NMHC: 2.0 x limit NO_x: +0.3/0.5 g/bhp-hr [2] PM: +0.04/0.05 g/bhp-hr
EOBD	Euro 5 (09/2009): CO: 1,900 mg/km NMHC: 250 mg/km NO_x: 300 mg/km PM: 50 mg/km [1] Euro 6-1 (09/2014): CO: 1,900 mg/km NMHC: 170 mg/km NO_x: 150 mg/km PM: 25 mg/km [1] Euro 6-2 (09/2017): CO: 1,900 mg/km NMHC: 170 mg/km NO_x: 90 mg/km PM: 12 mg/km [1]	Euro 5 (09/2009): CO: 1,900 mg/km NMHC: 320 mg/km NO_x: 540 mg/km PM: 50 mg/km Euro 6 interim (09/2009): CO: 1,900 mg/km NMHC: 320 mg/km NO_x: 240 mg/km PM: 50 mg/km Euro 6-1 (09/2014): CO: 1,750 mg/km NMHC: 290 mg/km NO_x: 180 mg/km PM: 25 mg/km Euro 6-2 (09/2017): CO: 1,750 mg/km NMHC: 290 mg/km NO_x: 140 mg/km PM: 12 mg/km	Euro IV (10/2005)/ Euro V (10/2008): NO_x: 7.0 g/kWh PM: 0.1 g/kWh NO_x control-system monitoring (since 11/2006): NO_x emission limit + 1.5 g/kWh Euro IV: (3.5+1.5) g/kWh Euro V: (2.0+1.5) g/kWh Euro VI-A (2013): NO_x: 1.5 g/kWh PM: 0.025 g/kWh (CI engine) Functional alternative for DPF monitor NO_x control system: SCR reagent NO_x: 0.9 g/kWh Euro VI-B (09/2014): SI engine Like Euro VI-A, but threshold value for CO: 7.5 g/kWh Euro VI-C (2016): NO_x: 1.2 g/kWh PM: 0.025 g/kWh (CI engine) CO: 7.5 g/kWh (SI engine) NO_x control system: SCR reagent NO_x: 0.46 g/kWh

[1] For gasoline direct injection.

[2] g/bhp-hr: grams per "brake horse power" times "hour" (brake horse power corresponds to the German unit PS).

fined for the particulate matter for gasoline engines from model year 2019 at a fixed value of 17.5 mg per mile. Further new requirements are, for example, the output of characteristics for assessing the use and activation of "Active off-cycle technologies" and fuel-consumption-specific quantities from model year 2019, improved monitoring of the crankcase-breather lines from model year 2023, and specification of a multitude of diagnostic requirements for components of hybrid vehicles.

Considerations are being given to extending the OBD requirements to CO_2 monitoring for future legislation.

Scope of application
The previously presented OBD regulations for the CARB and apply to all passenger vehicles with up to 12 seats as well as small commercial vehicles up to 14,000 lbs (6.35 t).

The current CARB OBD II legislation for California is at present also in force in some other US states. Furthermore, other US states are planning to adopt this legislation in future.

EPA OBD
Laws enforced by the EPA (Environmental Protection Agency) have been in force since 1994 in those US states which have not adopted the CARB legislation. The requirements of this diagnosis are essentially equivalent to the CARB legislation (OBD II). Within the framework of revising the Tier 3 emission legislation, from model year 2017 the EPA OBD requirements were adapted to the CARB OBD requirements. A CARB certificate is already recognized now by the EPA.

EOBD (European OBD)
On-board diagnostics attuned to European conditions is termed EOBD. EOBD has applied since January 2000 to passenger cars and light commercial vehicles with gasoline engines. The regulation has been valid for passenger cars and light commercial vehicles equipped with diesel engines since 2003, and for heavy commercial vehicles since 2005 (see OBD requirements for heavy commercial vehicles).

In 2007 and 2008 new EOBD requirements were adopted for gasoline and diesel passenger cars within the framework of the Euro 5 and Euro 6 emission and OBD legislation (Euro 5 emission level from September 2009; Euro 6 from September 2014).

A general new requirement for gasoline and diesel passenger cars is for checking the diagnostic frequency in everyday operation (In-Use Performance Ratio, IUPR) in accordance with the CARB OBD legislation (In-use Monitor Performance Ratio, IUMPR) as from Euro 5+ (September 2011).

EOBD Euro 5 and Euro 5+ requirements for diesel and gasoline engines
For gasoline engines, the introduction of Euro 5 as from September 2009 involved primarily a reduction in the OBD threshold values. In addition to a particulate-matter OBD threshold value (for direct-injection engines only), an NHMC OBD threshold value was introduced (non-methane hydrocarbons, instead of the previous HC). Direct functional OBD requirements result in the monitoring of the three-way catalytic converter for NMHC. Since September 2011 the Euro 5+ level has applied with unchanged OBD threshold values compared with Euro 5. A significant functional requirement with regard to EOBD is the additional monitoring of the three-way catalytic converter for NO_x.

For diesel passenger-car engines, Euro 5 involved a reduction of the OBD threshold values for particulate matter, CO and NO_x. In addition, there are extended requirements with regard to the monitoring of the exhaust-gas recirculation system (cooler) and above all with regard to the exhaust-gas treatment components. Here, monitoring of the SCR DeNOx system (dosing system and catalytic converter) is subject to very stringent requirements. Functional monitoring of the particulate filter is mandatory, irrespective of the untreated emissions.

EOBD Euro 6 requirements for diesel and gasoline vehicles
With Euro 6-1 from September 2014 and Euro 6-2 from September 2017 a further two-stage reduction of some OBD threshold values has been decided on (see Table 1). Furthermore, diesel systems are governed by tighter regulations for monitoring the oxidation catalytic converter and the NO_x exhaust-gas treatment system (NO_x storage catalytic converter or SCR catalytic converter with dosing system).

From September 2017 the NEDC is being replaced by the WLTC with Euro 6d-temp for the Type 1 emission test [8]. Here the emission limits as well as the OBD threshold values have not been adapted, but rather adopted unchanged. With regard to the OBD test the vehicle manufacturer can choose whether the die OBD threshold-value check is carried out on the basis of the NEDC or the WLTC. This right to choose will only exist until the end of 2021. From this point on, the OBD threshold values will exclusively be checked on the basis of the WLTC.

China OBD
In December 2016 the MEP (Ministry of Environmental Protection of the People's Republic of China) published a new law with significantly tightened emission and OBD requirements will come into force with regard to emissions in two stages (CN6a from 07/2020 and CN6b from 07/2023). The OBD requirements from 07/2020 will also apply unchanged to stage CN6b. While the previous Chinese legislation was geared very closely towards the European standard, the new CN6 law combines elements of the EU and US legislation and some country-specific new requirements. With regard to OBD, the requirements are based to the greatest possible extent on US requirements of the legislation from 2013, where some requirements have been removed or simplified and others have been enhanced. While the actual OBD requirements are geared towards the US standard, the European OBD threshold values for Euro 6-2 and the European test cycle WLTC have on the other hand been adopted.

Other countries
Some other countries have adopted different stages of the EU or the US OBD legislation (Russia, South Korea, India, Brazil, Australia).

OBD system requirements
The engine control unit must use suitable measures to monitor all on-board systems and components whose malfunction may cause a deterioration in exhaust-gas test specifications stipulated by law. A malfunction must be displayed to the driver by means of the malfunction indicator lamp (MIL) if a fault results in an excess in OBD threshold values.

OBD threshold values
US OBD II (CARB and EPA) prescribes thresholds that are defined based on emission limits. Accordingly, there are different permissible OBD threshold values for the various exhaust-gas categories that are applied during vehicle certification (e.g. LEV, ULEV, SULEV). The European EOBD regulations are based on absolute threshold values (Table 1).

Functional requirements
All exhaust-gas-related systems and components must, within the framework of On-Board Diagnostics (OBD) required by law, be monitored for malfunctions and for exceeded exhaust-gas threshold values.

Legislation demands the monitoring of electrical functions (short-circuit, line breaks), a plausibility check for sensors, and a function monitoring for actuators.

The pollutant concentration expected as the result of a component failure (can be measured in the exhaust-gas cycle) and the monitoring mode partly required by law determine the type of diagnostics. A simple functional test (black/white test) only checks system or component operability (e.g. swirl control valve opens or closes). The extensive functional test provides more detailed information about system operability and also identifies if necessary the quantitative influence of

defective components on emissions. As a result, the limits of adaptation must be monitored when monitoring adaptive fuel-injection functions (e.g. zero delivery calibration for a diesel engine, λ adaptation for a gasoline engine).

Diagnostic complexity has constantly increased as emission-control legislation has evolved.

Malfunction indicator lamp
The malfunction indicator lamp (MIL), also called the warning lamp, informs the driver that a component has malfunctioned. When a malfunction is detected, the CARB and the EPA stipulate that it must light up no later than after one driving cycle of its occurrence. In the area where EOBD applies, it must light up no later than the third driving schedule after the fault was detected.

If the malfunction disappears (e.g. a loose contact), the malfunction remains entered in the fault memory for 40 trips ("warm-up cycles"). The malfunction indicator lamp goes out after three fault-free driving schedules. The malfunction indicator lamp flashes for faults in the gasoline system if such a fault could cause damage to the catalytic converter (e.g. combustion misses).

Communication with the scan tool
The OBD legislation prescribes standardization of the fault-memory information and access to the information (connector, communication interface) compliant with the standard ISO 15031 and the corresponding SAE standards (e.g. SAE J1979, [2]). This permits a fault-memory readout over standardized, commercially available testers (scan tools).

Only diagnostics via CAN (ISO 15765 [4]) has been permitted for CARB since 2008 and for the EU since 2014.

Vehicle repair
Any workshop can use a scan tool to read out emission-related fault information from the control unit. This permits even non-franchised workshops to carry out repairs.

Manufacturers are obliged to provide the required tools and information (repair manuals on the Internet), for a suitable fee, to ensure that repairs can be carried out with the correct degree of expertise.

Switch-on conditions
The diagnostic functions are only executed if the physical switch-on conditions are fulfilled. These include, for example, torque thresholds, engine-temperature thresholds, and engine-speed thresholds or limits.

Inhibit conditions
Diagnostic functions and engine functions cannot always operate simultaneously. There are inhibit conditions that prohibit the performance of certain functions. For instance, tank ventilation (with evaporative-emissions control system) in the gasoline engine cannot function while catalytic-converter diagnosis is in operation. In the diesel engine the hot-film air-mass meter can only be monitored satisfactorily if the exhaust-gas recirculation valve is closed.

Temporary interruption of diagnostic functions
Diagnostic capabilities may only be disabled under certain conditions in order to prevent false diagnosis. Examples of such conditions are high elevation (low air pressure), during engine starting low ambient temperature, or low battery voltage.

Readiness codes
When the fault memory is checked, it is important to know that the diagnostic functions have run at least once. This can be checked by reading out the readiness codes over the diagnostics interface. These readiness codes are set for the most important monitored components on completion of the relevant diagnoses required by law.

Diagnostic System Management
The diagnostic capabilities for all components and systems checked must regularly run in driving mode, but also at least once during the exhaust-gas test cycle (e.g. FTP 75, NEDC). The Diagnostic System Management (DSM) can dynamically change the sequence for running the diagnostic functions depending on the driving condition. The objective here is to run the diagnostic functions frequently in everyday vehicle operation. Diagnostic System Management consists of the following components:
- Diagnostic Fault Path Management for storing fault states and associated ambient conditions (freeze frames)
- Diagnostic Function Scheduler for coordinating the engine and diagnostic functions
- Diagnostic Validator for deciding centrally when faults are detected whether they are the cause or a consequential fault. As well as central validation there are also systems with decentralized validation, i.e. validation is performed in the diagnostic function.

Vehicle recall
If vehicles fail to comply with OBD requirements by law, the authorities may demand the vehicle manufacturer to start a recall at their own cost.

OBD functions

Overview
Whereas EOBD only contains detailed monitoring specifications for individual components, the specific requirements in CARB OBD II are much more detailed. The list below shows the current state of the most significant CARB requirements (as at 2017) for gasoline-engined and diesel-engined vehicles. The requirements that are also described in detail in the EOBD legislation are marked by (E).

Gasoline and diesel system:
- Exhaust-gas recirculation system (E)
- Cold-starting emission-control system
- Crankcase ventilation
- Combustion misses/misfiring (E, for gasoline system only)
- Fuel system
- Variable valve timing
- Exhaust-gas sensors (λ oxygen sensors (E), NO_x sensors (E), particulate sensor)
- Engine cooling system
- Air-conditioning system (in the event of influence on emissions or on OBD)
- Other emission-related components and systems (E)
- In-use Monitor Performance Ratio (IUMPR) for checking the frequency of diagnostic functions in everyday operation (E).
- Requirements which apply to diesel/(gasoline) engines must, if the same technologies are used for the gasoline/(diesel) engine, be evaluated according to the diesel/(gasoline) requirements and the diagnostic concept of the authority/agency presented.

Gasoline system only:
- Secondary-air injection
- Three-way catalytic converter (E), heated catalytic converter
- Tank-leak diagnosis, with (E) at least electrical testing of the canister-purge valve
- Direct ozone-reduction system
- Cylinder-specific λ irregularity

Diesel system only:
- Oxidation-type catalytic converter (E)
- SCR DeNOx system (E)
- NO_x storage catalytic converter (E)
- Particulate filter (E)
- Fuel-injection system (rail-pressure control, injected fuel quantity, and injection timing)
- Cooler for exhaust-gas recirculation (E)
- Boost-pressure control
- Intercooler

Other emission-related components and systems refer to components and systems not mentioned in this list that, if they malfunction, may cause the exhaust-gas emissions to be increased (CARB-OBD II), the OBD threshold values to be exceeded (CARB OBD II and EOBD), or the diagnostic system to be negatively influenced (e.g. by inhibiting other diagnostic functions). Minimum values must be maintained with regard to the frequency of diagnostic functions.

Examples of OBD functions

Catalytic-converter diagnosis

Gasoline system

This diagnostic function monitors the conversion efficiency of the three-way catalytic converter. This is measured by the catalytic converter's oxygen retention capability. Monitoring is performed by observing the signals from the λ oxygen sensors in reaction to a specific alteration of the setpoint value of the λ closed-loop control.

Additionally, the NO_x storage capacity (catalytic-converter quality factor) must be assessed for the NO_x storage catalytic converter. For this purpose, the actual NO_x accumulator content resulting from consumption of the reduction agent during regeneration of the catalytic converter is compared with an expected value.

Diesel system

In the diesel system, carbon monoxide (CO) and unburned hydrocarbons (HC) are oxidized in the oxidation-type catalytic converter (pollutant minimization). Diagnostic functions are used to monitor the operation of the oxidation-type catalytic converter on the basis of the temperature differential before and after the catalytic converter (exothermy).

The storage and regeneration capacity of the NO_x storage catalytic converter is monitored. The monitoring functions run based on loading and regeneration models, and the measured regeneration duration. This requires the use of λ or NO_x sensors.

The SCR DeNOx catalytic converter is monitored using efficiency diagnostics. NO_x sensors located before and after the catalytic converter are required for this purpose. The components of the dosing system and the quantity and dosing of the reducing agent are monitored separately.

Tank-leak diagnosis

Gasoline system

Tank-leak diagnosis detects evaporation from the fuel system that may cause an increase in HC values, in particular. EOBD is limited to simply testing the electrical control circuit of the tank-pressure sensor and the canister-purge valve (evaporative-emissions control system). In the USA, on the other hand, it must be possible to detect leaks in the fuel system. There are two different methods of doing this.

The low-pressure method observes the tank pressure and first tests its operability by deliberately actuating the tank ventilation and carbon-canister check valves. A conclusion can then be drawn on the leak size using the time curve of the tank pressure – again by deliberately actuating the valves.

The overpressure method uses a diagnosis module with an integrated electrically powered vane pump that can be used to pump up the tank system. The flow from the pump is high when the tank is hermetically sealed. A conclusion can then be drawn on leak size by evaluating the flow from the pump.

Particulate-filter diagnosis
Diesel system
The diesel particulate filter is currently monitored for the most part for filter breakage, removal, or blockage. A differential-pressure sensor is used to measure the pressure differential (exhaust-gas backpressure downstream and upstream of the filter) at a specific volumetric flow. The measured value can be used to verify whether the filter is defective. An extended function uses load models to monitor the efficiency of the particulate filter.

Since model year 2010 the regeneration frequency has also had to be monitored. Since model year 2013 a particulate sensor has been used to monitor the particulate filter in response to tightened OBD requirements in the USA. The particulate sensor (from Bosch) operates according to the "collection principle", i.e. the soot collected over a specific distance driven is evaluated using a model for a threshold-value filter. If the collected soot matter, as a function of different parameters, exceeds a certain threshold, the particulate filter is detected as faulty. Combined particulate-filter faults (e.g. broken and melted filters) can also be detected with the particulate sensor.

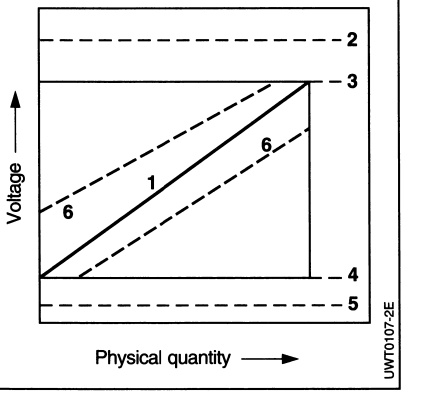
Exhaust-gas recirculation system diagnosis
Diesel system
In the exhaust-gas recirculation system (EGR) the regulator, the exhaust-gas recirculation valve, the exhaust-gas cooler, and other individual components are monitored.

Functional system monitoring is performed by air-mass regulators and position controllers. They check for permanent control variances. An excessively high or low EGR throughflow must be detected. System response ("slow response") must also be monitored.

The exhaust-gas recirculation valve itself is monitored for its electrical and functional operability.

The EGR cooler is monitored by means of additional temperature measurement after the cooler and with model values. This allows the cooler efficiency to be calculated.

Figure 2: Sensor monitoring
1 Sensor curve,
2 Upper threshold for "Signal Range Check",
3 Upper threshold for "Out of Range Check",
4 Lower threshold for "Out of Range Check",
5 Lower threshold for "Signal Range Check",
6 Plausibility check "Rationality Check".



Comprehensive components
On-board diagnostics requires that all sensors (e.g. air-mass meter, speed sensor, temperature sensors) and actuators (e.g. throttle valve, high-pressure pump, glow plugs) having either an impact on emissions or are being used to monitor other components or systems (and consequently may if necessary disable other diagnoses) have to be monitored.

Sensors monitor the following faults (Figure 2):
- Electrical faults, i.e. short-circuits and line breaks ("Signal Range Check").
- Range faults ("Out of Range Check"), i.e. undercutting or exceeding voltage limits set by the sensors' physical measurement range.
- Plausibility faults ("Rationality Check"); these are faults that are inherent in the components themselves (e.g. drift), or which may be caused by shunts, for instance. Monitoring is carried out by a plausibility check on the sensor signals, either by using a model or directly by other sensors.

Actuators must be monitored for electrical faults and – if technically possible – also for function. Functional monitoring means that, when a control command (setpoint value) is given, it is monitored by observing or measuring (e.g. by a position sensor) the system reaction (actual value) in a suitable way by using information from the system.

The actuators to be monitored include all output stages, the throttle valve, the exhaust-gas recirculation valve, the variable turbine geometry of the exhaust-gas turbocharger, the swirl flap, the injectors, the glow plugs (for diesel systems), the tank-ventilation system (for gasoline systems), and the active-charcoal check valve (for gasoline systems).

OBD requirements for heavy commercial vehicles

Europe

For commercial vehicles, the first stage of On-Board Diagnostics was introduced in the EU (EOBD) together with Euro IV (10/2005), the second stage together with Euro V (10/2008). A new OBD regulation came into force together with Euro VI in 2013.

Monitoring requirements in Stage 1
- Fuel-injection system: monitoring for electrical faults and for total failure.
- Engine components: monitoring of emission-related components for compliance with the OBD threshold value.
- Exhaust-gas treatment systems: monitoring for serious faults.

Additional requirements in Stage 2
- Exhaust-gas treatment systems: monitoring for compliance with the OBD threshold value.

Additional requirements

Since November 2006 it has been required to monitor the NO_x control systems for correct operation. The systems are monitored for their own emission limits, which are more stringent than the OBD threshold values.

SCR system

The aim is to ensure that the systems are supplied with the correct reagent (urea/water solution, the customary brand name is AdBlue). The availability of the reagent must be monitored by way of the tank fill level. To check the correct quality, it is necessary to monitor the NO_x emissions either with an exhaust-gas sensor or alternatively via a quality sensor. In the latter case, it is also necessary to monitor for correct reagent consumption.

Exhaust-gas recirculation system

The exhaust-gas recirculation system is monitored for correctly recirculated exhaust-gas mass flow and for deactivation of exhaust-gas recirculation.

NO_x storage catalytic converters
Exhaust-gas sensors are used to monitor NO_x emission.

Monitoring of NO_x control systems
Faults in NO_x control systems must be permanently stored (i.e. non-erasable) for 400 days (9,600 hours). Engine power must be throttled if the NO_x OBD threshold value is exceeded or if the urea tank is empty.

Euro VI
The OBD part of the Euro VI regulation is based on the Global Technical Regulation (GTR) "World Wide Harmonized OBD" (WWH OBD). In terms of structure this WWH OBD GTR corresponds to the Californian OBD legislation (passenger cars and commercial vehicles).

WWH OBD leaves opens which monitoring operations are actually selected in a national regulation (here Euro VI) which implements WWH OBD.

In addition, emission limits and OBD threshold values as well as the choice of test cycles are established by way of the national regulations.

Special aspects of WWH OBD involve the introduction of new fault storage and new scan-tool communication (ISO 27145 [5]).

Faults must be classified according to their severity with regard to emission deterioration. Emission-related faults can be distinguished by means of malfunction-indicator-lamp behavior and scan-tool communication. The following different categories are applied:
- A: Emissions above the OBD threshold value.
- B1: Emissions above or below the OBD threshold value.
- B2: Emissions below the OBD threshold value, but above the emission limit.
- C: Influence on emissions below the emission limit.

According to this principle all emission-related faults are output, even those which exert a very small influence.

Data for Euro VI
- Drastic reduction of the emission limits and OBD threshold values for NO_x and particulate matter compared with Euro V.
- Emission limits to be introduced for NH_3 and the number of particulates.
- Use of the harmonized WHSC and WHTC test cycles.
- The OBD demonstration is performed with a double WHTC hot start part.
- Introduction of a facility for checking system conformity with regard to emissions in the field by means of random measurements with portable emission measurement systems (PEMS).
- Checking of the diagnostic frequency of OBD monitoring in everyday operation (In-Use Monitoring, IUMPR).

Euro VI A
- For new type approvals mandatory since 31 December 2012.
- Valid up to 31 August 2015.
- Strict OBD threshold values for NO_x and particulate matter (only for engines with auto-ignition). For particulate-filter monitoring a functional, non-emission-correlated diagnosis is possible as an alternative to OBD threshold-value diagnosis.

Euro VI B
- For new type approval mandatory as from 1 September 2014.
- Valid up to 31 December 2016.
- Affects only engines with externally supplied ignition.
- Introduction of an OBD threshold value for CO.

Euro VI C
- For new type approval mandatory as from 31 December 2015.
- Changes in respect of Euro VI A and Euro VI B: Tightened NO_x OBD threshold value and tightening of the NO_x control-system requirements for SCR reagent quality and consumption monitoring. Monitoring is performed with regard to the long-term drift behavior of the fuel injectors. Monitoring of the OBD diagnostic-frequency rate is mandatory.

Euro VI D
- For new type approval mandatory as from 1 September 2018.
- No change with regard to OBD.

Required diagnoses from WWH OBD
These diagnoses are compulsory for particulate filter, SCR catalytic converter, NO_x storage catalytic converter, oxidation catalytic converter, exhaust-gas recirculation, fuel-injection system, charge-air pressure system, variable valve control, cooling system, exhaust-gas sensors, idle control system, and components.

Required diagnoses outside the WWH OBD scope
For the particulate filter, the exhaust-gas recirculation system and the charge-air pressure control system no exceptions are permitted in monitoring for specific diagnoses. The relevant faults must not be defined as Class C.

Furthermore, as from Euro VI C the monitoring of possible component-damaging effects of a long-term drift of fuel injectors is required.

Fundamentally the definition of the non-emission-correlated "performance monitors" in WWH OBD has been changed. In Euro VI these diagnoses must be demonstrated with emission correlation for the first certification of an engine from an engine family.

Gas engines
Gas engines are covered by specific monitoring requirements with regard to compliance with the λ setpoint value, to NO_x and CO conversion of the three-way catalytic converter, and to the λ sensor. Furthermore, catalytic-converter-damaging combustion-miss detection is required.

Requirements with regard to NO_x control systems
For SCR systems, monitoring of the reagent tank fill level, of the reagent quality, of reagent consumption, and of dosage interruption is required.

For exhaust-gas recirculation systems, monitoring of the exhaust-gas recirculation valve is required.

Furthermore, all the NO_x control systems must be monitored for system deactivation by tampering.

Detected faults in the NO_x control system result in a reduction in stages of vehicle drivability. Torque limitation as the first stage is followed by vehicle-speed limitation to crawling speed as the second stage.

USA
CARB, model year 2007
In California "Engine Manufacturer Diagnostics" (EMD) has been called on for heavy commercial vehicles since model year 2007. This can be viewed as a precursor to an OBD regulation. EMD calls for the monitoring of all components and of exhaust-gas recirculation.

The requirements are not reflected in a separate emission limit. Furthermore, no standardized scan-tool communication is required.

Model year 2010 and following
With model year 2010 an OBD system as for passenger-car OBD II was introduced. The technical requirements are at the same level as the respective requirements for passenger cars.

Differences ensue due to the fact that commercial vehicles are subject to engine certification. All the emission limits and OBD threshold values apply here to engine cycles. The absolutely applicable values are scaled with the work performed in the cycle.

In contrast to passenger-car LEV III, a new emission regulation (introducing an NO_x and NMHC aggregate limit value) is not planned for commercial vehicles. The OBD threshold values for NO_x and NMHC thus remain unchanged, i.e. separate, for commercial vehicles.

Time frame for introducing OBD requirements
Model year 2010
A performance variant of a manufacturer's top-selling engine family must be equipped with an OBD system. The other performance variants of this engine family are covered by a simplified certification procedure.

Model year 2013
An engine family of a manufacturer must be equipped in all the performance variants with an OBD system. Furthermore, an OBS system is required in each case for a performance variant in every engine family. The other performance variants of these engine families are covered by a simplified certification procedure.

Model year 2016
All the engine families of a manufacturer must have an OBD system in all the performance variants.

Model year 2018
Engines powered by alternative fuels (e.g. gas) are subject to the OBD requirements.

Japan
Japan has had its own OBD regulation for commercial vehicles in force since 2004. The requirements are comparable in terms of content with EMD in California for model year 2007.

China
China V has applied in China since 2017; this conforms with regard to OBD to Euro V.

According to a draft China VI will be introduced from 2020; this is in fact based on Euro VI, but in contrast is to impose additional requirements that relate to OBD:
- Additional demonstration of the OBD diagnoses at the vehicle on a special commercial-vehicle chassis-dynamometer test cycle C-WTVC.
- Monitoring of the injected fuel quantity with regard to the OBD threshold values.
- Remote transmission of live data which are determined via the OBD system by way of remote data transmission to a server.
- So-called "Permanent DTC" fault storage similar to CARB OBD, adapted to WWH OBD fault storage.
- Special temperature monitoring for vanadium catalytic converters.
- Use of a warning and prompt system, as there is for Euro VI for NO_x control-system requirements, for OBD faults too. This relates to the OBD faults for a loss of efficiency of the most important exhaust-gas treatment systems DPF (diesel particulate filter), SCR catalytic converter (selective catalytic reduction), NO_x storage catalytic converter, and three-way catalytic converter.
- Specifically for gas engines: monitoring of crankcase ventilation.

Other countries
Further countries have in the meantime introduced OBD for commercial vehicles. These include India, Korea, Australia, Brazil, and Russia. These countries have adopted EU regulations in this regard (Euro IV, Euro V or Euro VI (Korea)).

On-Board Diagnostics for motorcycles

While there have hitherto been no OBD requirements for motorcycles in the USA, OBD is being introduced in three stages in the EU. Emission stage Euro 4 as from 2016 saw the introduction of OBD I, which requires the monitoring of electrical faults for emission-relevant sensors and actuators. Which diagnoses of electrical faults are required is dependent on the component. The additional requirements such as fault detection and fault storage, MIL activation, scan-tool communication, etc. conform to the EOBD passenger-car requirements.

With emission stage Euro 5 as from 2020 OBD I will be extended to include plausibility checks for sensors and function checks for actuators. The diagnostic range is dependent on the component. It will furthermore be necessary to monitor for faults which lead to a reduction of the maximum torque by more than 10%. At the same time OBD II requirements as for EOBD apply as well. Catalytic-converter diagnosis is excluded. The detection of misfires on motorcycles is covered by engine-speed load ranges. The functional IUPR requirements must be satisfied, but without prescribed minimum ratios. From 2024 reduced OBD threshold values will reply; furthermore, catalytic-converter diagnosis is prescribed and a minimum ratio of 0.1 will apply to all diagnoses.

The first OBD I stage was carried over to UN GTR 18, which alternatively also provides for reduced scopes. The carryover of OBD I stage 2 and of OBD II to a UN GTR is planned for 2018. India is expected to introduce EU OBD I from 4/2020 and EU OBD II from 4/2024. It is unclear whether India will adopt the separation of OBD II into two stages as in the EU.

Japan introduced its own OBD I with a similar scope to EU OBD I in 2016. A further stage OBD II is to be based on the previously mentioned UN GTR OBD II.

China has its own simplified variant of the EU OBD I legislation, valid from 7/2018.

ECU diagnostics and Service Information System

Function

The function of diagnostics in the workshop is to locate quickly and reliably the smallest replaceable defective unit. In modern vehicles it is essential to use a generally PC-based diagnostic tester. Diagnostics in the workshop utilizes the results of diagnostics conducting in driving mode (fault-memory entries), employing special workshop diagnostic modules in the vehicle ECU or diagnostic tester and additional test and measuring equipment. These diagnostic options are integrated in the diagnostic tester in prompted troubleshooting.

Prompted troubleshooting

The main element of workshop diagnostics is the prompted troubleshooting procedure. The workshop employee is guided, starting out from the fault pattern or from the fault-memory entry, through the relevant diagnostic steps. The choice and order of the diagnostic steps is dynamic and depends on the result of the previous diagnostic step. The diagnostic results contain the use of test equipment, additional sensors or workshop diagnostic modules.

Symptoms

Faulty vehicle behavior can be either perceived directly by the driver or documented by way of a fault-memory entry. At the beginning of fault diagnosis the workshop employee must identify the existing symptom as the starting point for prompted troubleshooting.

Fault-memory entries

All the faults that occur during driving are stored in the fault memory together with defined ambient conditions prevailing at the time of their occurrence, and can be read out via an interface protocol. This protocol is based on one of the established standards and is generally extended to include manufacturer-specific components. The fault memory can also be deleted with the diagnostic tester.

Additional test equipment and sensors

The diagnostic options in the workshop are expanded by using additional sensors (e.g. clamp-on current pickup, clamp-on pressure pickup) or test equipment (e.g. Bosch vehicle-system analyzer). In the event of a fault detected in the workshop, the equipment is adapted to the vehicle. The measurement results are generally evaluated by the diagnostic tester.

Workshop diagnostic modules

All the workshop diagnostic modules can be used only when a diagnostic tester is connected and generally only when the vehicle is stationary. The operating conditions are monitored in the ECU. These diagnostic modules integrated in the ECU, after being started by the diagnostic tester, run completely autarkically in the ECU and on completion signal the result back to the diagnostic tester. ECU-based workshop diagnostic modules differ from simple actuator tests with acoustic feedback in that they can place the vehicle to be diagnosed in the workshop into as a rule no-load operating points, excite actuators depending on the boundary conditions set, and independently evaluate the result via sensor values with an evaluator logic circuit. Examples of such modules are the high-pressure test as system test for the diesel injection system (Figure 1) and the run-up test as component test for the diesel injectors (Figure 2).

Tester-based diagnostic modules

The functional sequence and evaluation are executed in the case of tester-based diagnostic modules in the diagnostic tester, where the measurement data used for evaluation are determined with the aid of the ECU by sensors in the vehicle or by additional test sensors.

The performance of the tester-based diagnostic modules is dependent on the released interfaces in the engine ECU and the data transfer between engine ECU and tester. On the other hand, high flexibility

Figure 1: High-pressure test for the diesel injection system

Detection of leaks in the high-pressure system and efficiency of high-pressure generation.
Steps:
Start test – Increase setpoint pressure – Measure pressure build-up time – Reduce setpoint pressure – Measure pressure reduction time – Variation of engine speed and pressure – Measure pressure reduction time with engine stopped at end of test – Diagnostic results.

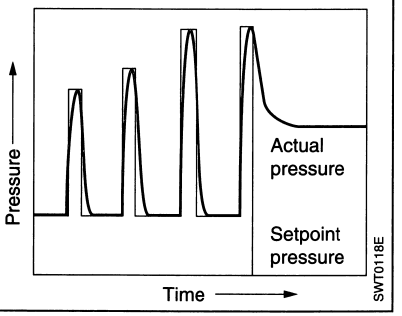
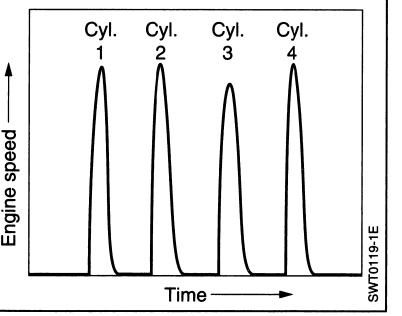


Figure 2: Run-up test for the diesel injectors

Detection of deviations of injected fuel quantity of individual injectors.
Steps:
Start test – Cutout of individual cylinder – Quantity jump – Measure maximum speed – Repeat with further cylinders – Diagnostic result.



is also possible after series introduction since new diagnostic modules can also be used thanks to a simple update of the tester software. Selective diagnostics is therefore possible when field problems occur.

Further contents of fault-finding instructions

The fault-finding instructions also include function descriptions, electrical terminal diagrams, installation positions for components, read-outs of actual values, and actuator diagnostics.

References

[1] ISO 15031: Road vehicles – Communication between vehicle and external equipment for emissions-related diagnostics (2011).
[2] SAE J 1979: E/E Diagnostic Test Modes (2012).
[3] SAE J 1939: Serial Control and Communications Heavy Duty Vehicle Network – Top Level Document (2012).
[4] ISO 15765: Road vehicles – Diagnostics over Controller Area Network (2011).
[5] ISO 27145: Road vehicles – Implementation of World-Wide Harmonized On-Board Diagnostics (WWH-OBD) communication requirements (2012).
[6] OBD II regulation, section 1968.2 of title 13, California Code of Regulations, different approved OAL versions.
[7] UN/ECE Regulation No. 83, Revision 5: Uniform provisions concerning the approval of vehicles with regard to the emission of pollutants according to engine fuel requirements.
[8] (WLTP) Regulations (EU) 2017/1151 and 2017/1347.